

## *SPECIAL ISSUE: TEACHING CASES*

In this issue:

- 4. Student Guide: Super Saturday Series (S3): Dashboard in Power View**  
Nannette P. Napier, Georgia Gwinnett College
- 11. ACS: Bringing Business Intelligence and Analytics to a Massive Multiplayer Online Gaming Company**  
Nicholas Christian, Middle Tennessee State University  
Innocent Akujobi, Middle Tennessee State University  
Munmun Saha, Middle Tennessee State University  
Melinda Korzaan, Middle Tennessee State University
- 18. Dragon Air: A Database Design Case**  
Donald Wynn Jr., University of Dayton
- 25. System Development and Data Modeling for Stevens' Wholesale Health Supplies**  
Dana Schwieger, Southeast Missouri State University  
Ziping Liu, Southeast Missouri State
- 34. Formula One – a database project from start to finish**  
Anthony Serapiglia, St. Vincent College

The **Information Systems Education Journal (ISEDJ)** is a double-blind peer-reviewed academic journal published by **ISCAP** (Information Systems and Computing Academic Professionals). Publishing frequency is six times per year. The first year of publication was 2003.

ISEDJ is published online (<http://isedj.org>). Our sister publication, the Proceedings of EDSIGCON (<http://www.edsigcon.org>) features all papers, panels, workshops, and presentations from the conference.

The journal acceptance review process involves a minimum of three double-blind peer reviews, where both the reviewer is not aware of the identities of the authors and the authors are not aware of the identities of the reviewers. The initial reviews happen before the EDSIGCON conference. At that point papers are divided into award papers (top 15%), other journal papers (top 30%), unsettled papers, and non-journal papers. The unsettled papers are subjected to a second round of blind peer review to establish whether they will be accepted to the journal or not. Those papers that are deemed of sufficient quality are accepted for publication in the ISEDJ journal. Currently the target acceptance rate for the journal is under 40%.

Information Systems Education Journal is pleased to be listed in the Cabell's Directory of Publishing Opportunities in Educational Technology and Library Science, in both the electronic and printed editions. Questions should be addressed to the editor at [editor@isedj.org](mailto:editor@isedj.org) or the publisher at [publisher@isedj.org](mailto:publisher@isedj.org). Special thanks to members of AITP-EDSIG who perform the editorial and review processes for ISEDJ.

### **2018 AITP Education Special Interest Group (EDSIG) Board of Directors**

Leslie J. Waguespack Jr  
Bentley University  
President

Jeffry Babb  
West Texas A&M University  
Vice President

Scott Hunsinger  
Appalachian State Univ  
Past President (2014-2016)

Amjad Abdullat  
West Texas A&M University  
Director

Meg Fryling  
Siena College  
Director

Li-Jen Lester  
Sam Houston State Univ  
Director

Lionel Mew  
University of Richmond  
Director

Rachida Parks  
Quinnipiac University  
Director

Anthony Serapiglia  
St. Vincent College  
Director

Jason Sharp  
Tarleton State University  
Director

Peter Wu  
Robert Morris University  
Director

Lee Freeman  
Univ. of Michigan - Dearborn  
JISE Editor

Copyright © 2018 by Information Systems and Computing Academic Professionals (ISCAP). Permission to make digital or hard copies of all or part of this journal for personal or classroom use is granted without fee provided that the copies are not made or distributed for profit or commercial use. All copies must bear this notice and full citation. Permission from the Editor is required to post to servers, redistribute to lists, or utilize in a for-profit or commercial use. Permission requests should be sent to Jeffry Babb, Editor, [editor@isedj.org](mailto:editor@isedj.org).

# INFORMATION SYSTEMS EDUCATION JOURNAL

## Editors

**Jeffrey Babb**  
Senior Editor  
West Texas A&M University

**Anthony Serapiglia**  
Teaching Cases Co-Editor  
St. Vincent College

**Guido Lang**  
Associate Editor  
Quinnipiac University

**Thomas Janicki**  
Publisher  
U of North Carolina Wilmington

**Paul Witman**  
Teaching Cases Co-Editor  
California Lutheran University

**Muhammed Miah**  
Associate Editor  
Southern Univ at New Orleans

**Donald Colton**  
Emeritus Editor  
Brigham Young Univ. Hawaii

**James Pomykalski**  
Associate Editor  
Susquehanna University

**Jason Sharp**  
Associate Editor  
Tarleton State University

## 2018 ISEDJ Editorial Board

Nita Brooks  
Middle Tennessee State Univ

Wendy Ceccucci  
Quinnipiac University

Ulku Clark  
U of North Carolina Wilmington

Jamie Cotler  
Siena College

Christopher Davis  
U of South Florida St Petersburg

Gerald DeHondt II

Mark Frydenberg  
Bentley University

Meg Fryling  
Siena College

Biswadip Ghosh  
Metropolitan State U of Denver

David Gomillion  
Northern Michigan University

Janet Helwig  
Dominican University

Scott Hunsinger  
Appalachian State University

Mark Jones  
Lock Haven University

James Lawler  
Pace University

Li-Jen Lester  
Sam Houston State University

Michelle Louch  
Duquesne University

Lionel Mew  
University of Richmond

George Nezelek  
Univ of Wisconsin Milwaukee

Rachida Parks  
Quinnipiac University

Alan Peslak  
Penn State University

Doncho Petkov  
Eastern Connecticut State Univ

Samuel Sambasivam  
Azusa Pacific University

Karthikeyan Umapathy  
University of North Florida

Leslie Waguespack  
Bentley University

Bruce White  
Quinnipiac University

Peter Y. Wu  
Robert Morris University

## Teaching Case

# Student Guide: Super Saturday Series (S3) Dashboard in Power View

Nannette P. Napier  
nnapier@ggc.edu  
School of Business Georgia  
Gwinnett College Lawrenceville, GA 30043, USA

### Abstract

Dashboards allow managers to interact with data, asking “what if” questions and looking for patterns. This topic is typically mentioned in information systems courses; however, many students will not have experience with developing dashboards or using them to answer business questions. The purpose of this assignment is to provide students with hands-on experience creating a dashboard using Excel’s free Power View add-in. The exercise is based upon a real case. Students are provided anonymized registration data for participants at the Super Saturday Series (S3), an outreach event for middle and high school girls. The dashboard is designed for the S3 Directors to answer questions such as: How much money was collected? Which day and time of day was the most popular for registrations? How many participants attended each date? Which school systems had the highest number of attendees? Lowest number of attendees? Step-by-step instructions demonstrate how to add different types of visualizations, including line charts, bar charts, and tables. Students will customize their dashboard by selecting a theme and inserting an image. After completing the dashboard, students will use it to answer business questions. This assignment is suitable for undergraduate students taking a computer literacy, introductory information systems, or data analytics course. Although some basic proficiency in Excel is required, students do not need any prior experience in Power View to complete this assignment.

**Keywords:** Excel, Computer literacy, Data analytics, introduction to information systems

## 1. BACKGROUND INFORMATION

Super Saturday Series (S3, [www.ggc.edu/s3](http://www.ggc.edu/s3)) is a fun, hands-on workshop series that exposes middle and high school girls to technology and science in a friendly college setting. S3 utilizes interactive games and experiments to pique girls’ interest in technology and science. S3 also provide girls with an exciting opportunity to interact with positive, professional role models. Our community, comprised of students and volunteers, promotes learning in an engaging way in order to lay a positive science and technology foundation that attendees’ current education may lack.

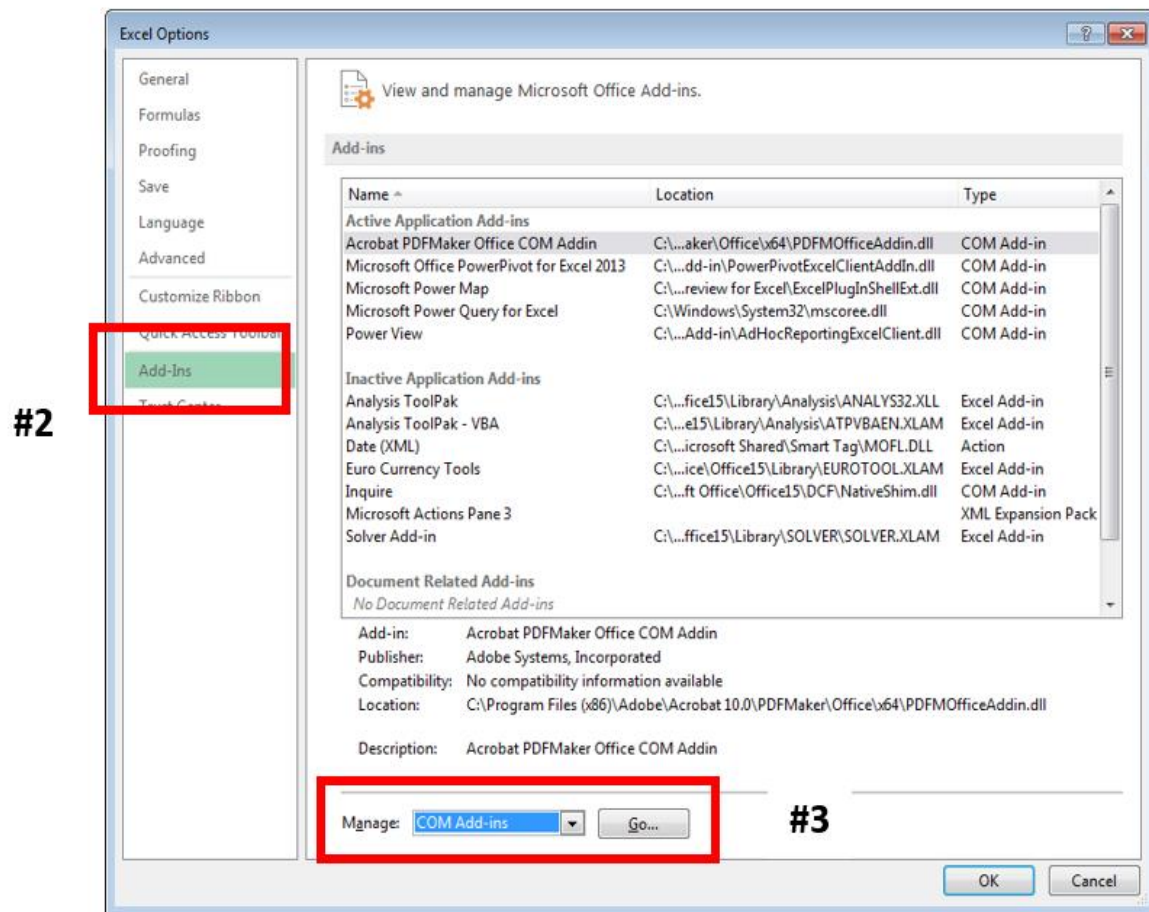
Parents enroll students in S3 using an online form and enter personal information such as participant’s name, email address, grade level,

gender, ethnicity, school, and desired registration date. The online system documents the date and time that parents register, and S3 staff keep track of payments made.

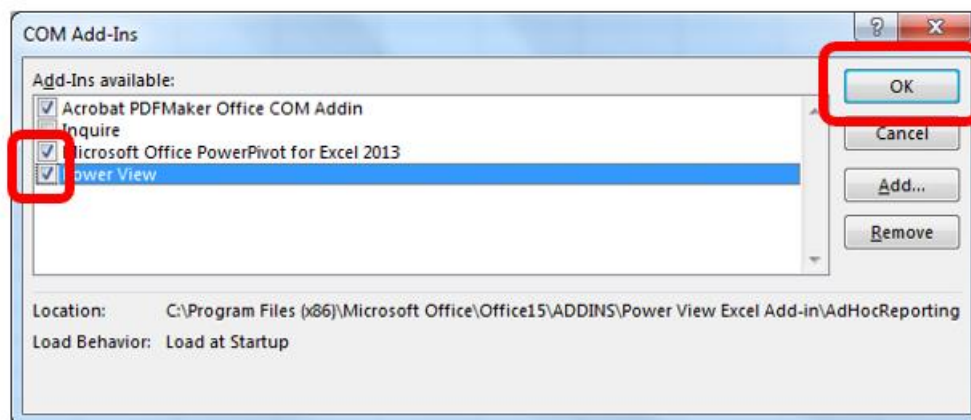
The S3 Directors have provided you an anonymized version of registration data for the February 21st and February 28th sessions. To assist them, you will create a dashboard within Excel using Microsoft’s Power View add-in.

## 2. SETUP IN EXCEL

- 1) Open Excel 2013 or 2016.
- 2) Go to File → Options → Add-Ins
- 3) In the Manage box, click COM Add-ins → Go.

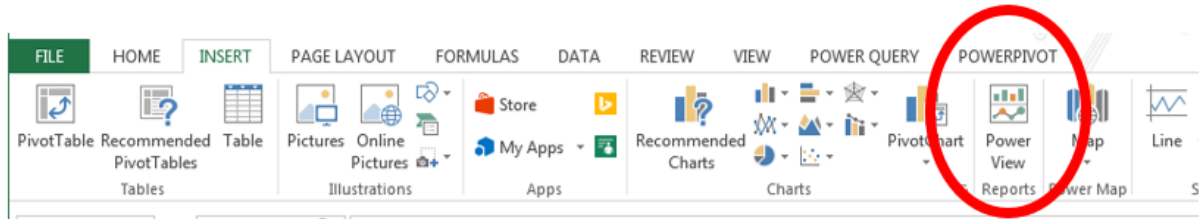


**Figure 2.1: Managing Add-Ins**

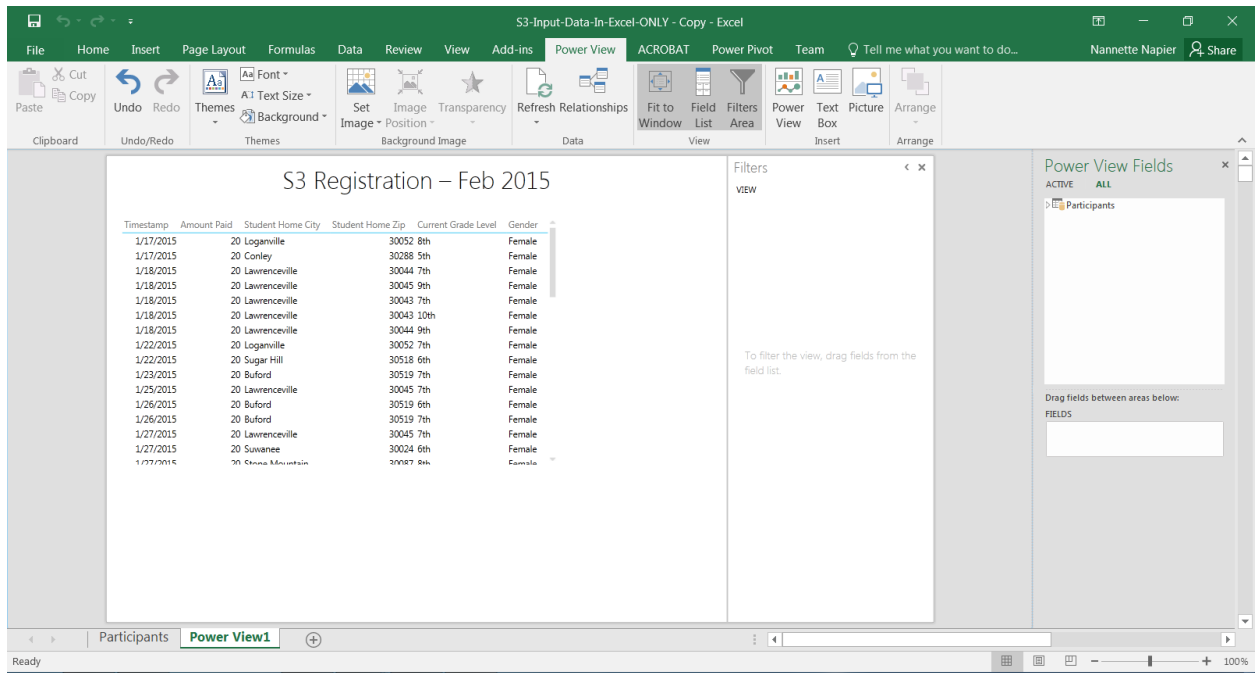


**Figure 2.2: Selecting Add-ins**

- 4) On the screen that appears, check the boxes for
  - Microsoft Office Power Pivot in Microsoft Excel 2013/2016
  - Power View
  - Then, Select OK
- 5) In Excel 2013, Power View should automatically appear on the Excel ribbon under the Insert tab after completing these steps.



**Figure 2.3 Excel Insert Tab with Power View**



**Figure 3.1 Power View with Updated Title**

In Excel 2016, you should follow these instructions to customize the ribbon to make Power View appear:

<https://support.office.com/en-us/article/Turn-on-Power-View-in-Excel-2016-for-Windows-f8fc21a6-08fc-407a-8a91-643fa848729a?ui=en-US&rs=en-US&ad=US>.

6) Close Excel

### 3. GETTING STARTED

- 1) Download S3-Input-Data.xlsx and save to a known location.
- 2) Open the file in Excel. You should see the *Participants* worksheet which contains registration data organized as an Excel table.
  - Each row provides information about one participant in the S3 program. As you can see, there are 101 rows

indicating that 101 girls registered as participants across the two dates: Feb 21 and Feb 28.

- Personal information captured (such as name, home address, and email) has been removed.
- 3) With your cursor inside the Excel table, open Power View (Insert tab → Power View). Please be patient as it may take a minute or two to load.
  - 4) When the Power View screen comes up, replace "Click here to add a title" with "S3 Registration – Feb 2015." So far, your screen should look like this:
  - 5) Delete the default table that was provided by selecting the outside edge of the table. Then, click **Delete** button.
  - 6) Save your work.

#### 4. ETHNICITY TABLE

- 1) Create a new visualization by selecting the following two fields under *Participants* as shown in Figure 4.1:
  - Ethnic Origin
  - Amount Paid

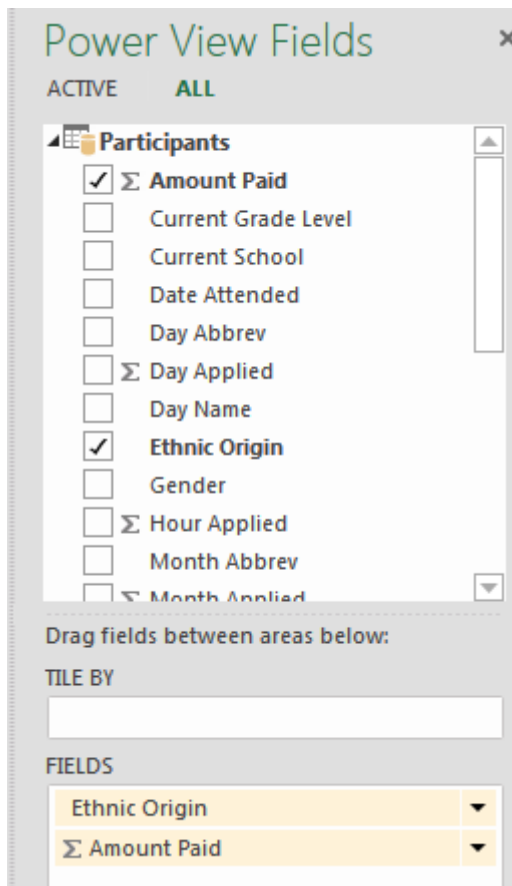


Figure 4.1 Fields for Ethnicity Table

- 2) Change the format of the values in Amount Paid by doing the following (Figure 4.2):
  - Select one of the values under the "Amount Paid" column.
  - On the Design tab in the Number group, change the format from General to Currency.
- 3) Click on the "Amount Paid" heading to sort the data. The data should appear in descending order by amount paid.
- 4) Above the chart, insert a Text Box to use as a title: "Participants by Ethnicity". *HINT:* The Text Box is on the Power View tab in the Insert group.
- 5) Resize the textbox as needed

Participants by Ethnicity	
Ethnic Origin	Amount Paid ▼
African American/Black	\$880.00
Other	\$410.00
White	\$360.00
Hispanic	\$140.00
Prefer not to answer	\$140.00
<b>Total</b>	<b>\$1,930.00</b>

Figure 4.3 Final Ethnicity Table

<i>Before:</i>		<i>After:</i>	
Ethnic Origin	Amount Paid	Ethnic Origin	Amount Paid
African American/Black	880	African American/Black	\$880.00
Hispanic	140	Hispanic	\$140.00
Other	410	Other	\$410.00
Prefer not to answer	140	Prefer not to answer	\$140.00
White	360	White	\$360.00
<b>Total</b>	<b>1930</b>	<b>Total</b>	<b>\$1,930.00</b>

Figure 4.2 Impact of Format Change

## 5. COLUMN CHART

- 1) Click on a blank part of the existing Power View canvas to indicate you want to create a new visualization.
- 2) Select the following two fields under *Participants* as shown in Figure 5.1:
  - Short Date Attended
  - Timestamp

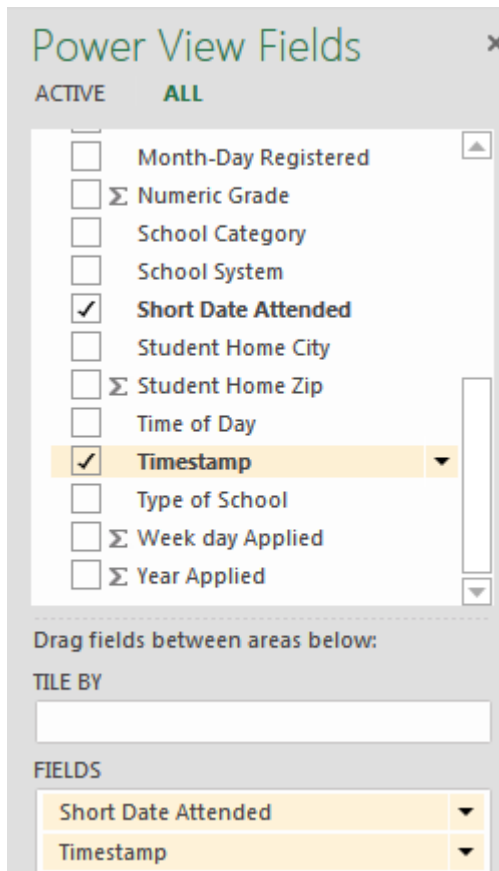


Figure 5.1 Fields for Column Chart

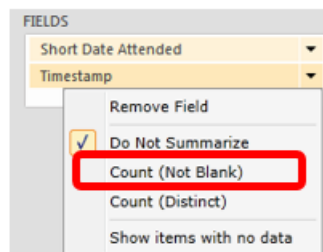


Figure 5.2 Change to Count Timestamp

- 3) The table created shows each individual timestamp. Instead, we want to show the Count of each timestamp (which represents

the count of participants). To make this change:

- Under FIELDS, click on Timestamp
  - Select "Count (Not Blank)"
- 4) Both the field description and actual visualization should change (as shown in Figure 5.3 below).

FIELDS	
Short Date Attended	
# Count of Timestamp	

Short Date Attended	Count of Timestamp
2/21/2017	40
2/28/2017	61
<b>Total</b>	<b>101</b>

Figure 5.3 Attendance Count by Date

- 5) Switch this visualization to a Stacked Column Chart. *HINT:* Design tab → Column Chart → Stacked Column Chart
- 6) Remove the title from this visualization  
*HINT:* Layout tab → Title → None
- 7) Show Data Labels  
*HINT:* Layout tab → Data Labels → Show
- 8) Resize as desired

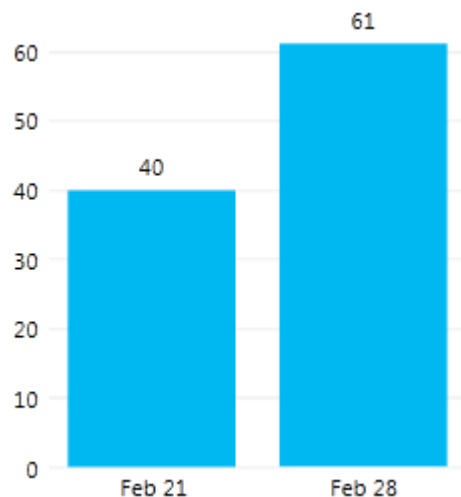
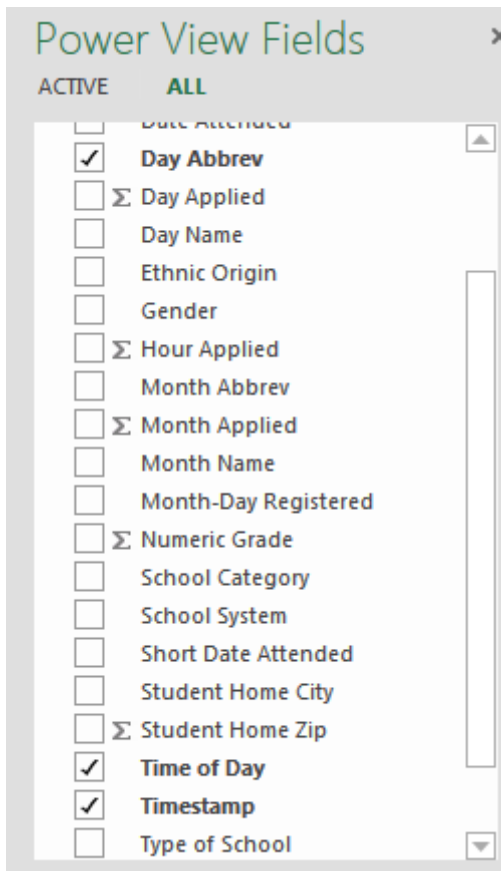


Figure 5.4 Final Column Chart

## 6. BAR CHART

- 1) Click on a blank part of the Power View canvas to indicate you want to create a new visualization.
- 2) Select the following two fields under *Participants*
  - Day Abbrev
  - Time of Day
  - Timestamp

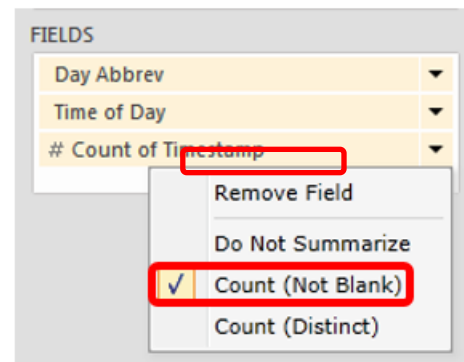




**Figure 6.1 Fields for Bar Chart**

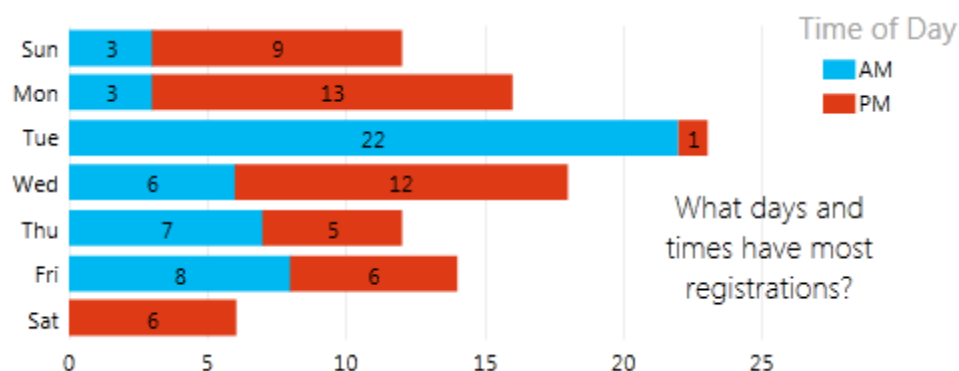
- 3) As before, change Timestamp to Count of Timestamp

- 4) Switch this visualization to a Stacked Bar Chart
- 5) Remove the title from this visualization
- 6) Show Legend at Right
- 7) Show Data Labels
- 8) Insert a Text Box to use as a title: "What days and times have most registrations?"



**Figure 6.2 Change to Count of Timestamp**

- 9) Switch this visualization to a Stacked Bar Chart
- 10) Remove the title from this visualization
- 11) Show Legend at Right
- 12) Show Data Labels
- 13) Insert a Text Box to use as a title: "What days and times have most registrations?"  
Feel free to resize and move the visualization on the canvas



**Figure 6.3 Final Bar Chart**

## 7. MATRIX TABLE

- 1) Create a new visualization using the following fields
  - a. School Category
  - b. School System
  - c. Timestamp
- 2) Change Timestamp → Count of Timestamp
- 3) Switch the visualization to Table → Matrix
- 4) Resize and move so there is enough room for the data about School systems and Count of Timestamp to appear on the dashboard

## 8. LINE CHART

- 1) Create a new visualization using the following fields
  - a. Month-Day Registered
  - b. Timestamp
- 2) Change Timestamp → Count of Timestamp
- 3) Switch the visualization to Line Chart
- 4) Remove the title

- 5) Insert text that can be used as a title: "Apps per Day"

## 9. FINISHING TOUCHES

- 1) Arrange all the visualizations within the Dashboard. *It will be a little crowded, but you can always use the "pop out" accelerator to see a particular visualization full screen.*
- 2) Insert picture of your school logo onto the dashboard [*HINT*: Power View tab → Picture]
- 3) Change the theme of the dashboard [*HINT*: Power View tab → Themes]
- 4) Change the background to something of your choice [*HINT*: Power View tab → Background]
- 5) Rename the worksheet from "Power View 1" to "S3 Dashboard"

When done, check your work. Save your file. Then, submit.

## *Teaching Case*

# ACS: Bringing Business Intelligence and Analytics to a Massive Multiplayer Online Gaming Company

Nicholas Christian  
ndc2b@mtmail.mtsu.edu

Innocent Akujobi  
ioa2b@mtmail.mtsu.edu

Munmun Saha  
ms9d@mtmail.mtsu.edu

Melinda Korzaan  
melinda.korzaan@mtsu.edu

Jennings A. Jones College of Business  
Middle Tennessee State University  
Murfreesboro, TN

### **Abstract**

Advanced Competition Simulations (ACS), a massive multiplayer online (MMO) strategy gaming company, has petabytes of customer data at their data center in Brentwood, Tennessee that could assist executives with decision making, but they currently do not have business intelligence software or a business intelligence and analytics (BIA) team. As is the case with most companies that want a new capability, they must decide if they want to hire a team internally or outsource this to a vendor. This decision was not taken lightly by ACS, but ultimately the business intelligence and analytics team was formed. Once the team was formed, the project to bring BIA to ACS gets off to a rough start when the team does not agree with the project manager choice. There are further setbacks when the data needs to be cleaned and the vendor refuses to help. Finally, the CEO of ACS questions his decision to rush the project when their data center has power failures during the data migration phase and he is reminded that the BIA project was moved ahead of the disaster recovery project. Bringing business intelligence and analytics to ACS proves to be much more difficult than they anticipated.

**Keywords:** IT Priorities, Business Intelligence and Analytics (BIA), Outsourcing Decision, Data Cleansing, Disaster Recovery Planning.

### **1. INTRODUCTION**

Jill walks past the seventy-inch interactive project smart board one Monday morning and realizes that Rich has added a new project to the top of the project priorities list, and this immediately infuriates her. She rushes down the hall to Rich's

office and shuts the door not too gently behind her. Before Rich has a chance to say anything, Jill practically yells, "Are you kidding me?" "What is this BIA project on the board and why did you put it above all of my IT projects?" Rich has a habit of rearranging the project board but this time he may have gone too far. He attempted to

argue that Business Intelligence and Analytics (BIA), which includes implementing data visualization, would help them make better business decisions about their primary focus: the new app that they were developing. But Jill was not buying it. Rich, while motioning for Jill to lower her voice, made the comment, "this project needs to go first and if you just trust me, you will see why. If you remember, you trusted my crazy idea once before and look where it got us." After a lengthy discussion, Jill began to see why Rich wanted to move forward with the BIA project, but she still did not agree with postponing their other IT projects. As she was walking out of his office, she made the comment, "I cannot believe you are going forward with this when our data center is incomplete and vulnerable to instability. We should at least finish the disaster recovery project first."

Jill currently has five of the top ten projects on the priorities list, but the projects are only at the top because Rich keeps skipping them.

## 2. ACS BACKGROUND

In April of 2014, Rich and Jill Amber were enjoying their first year of marriage in Silicon Valley, California when Rich had an awesome idea for a mobile massive multiplayer online (MMO) strategy gaming app. Rich and Jill were both developers for a small software company at the time, so they decided to write the app in their free time. Being avid gamers and experienced developers allowed them to create a game with all the features that they had always wished were in an app. This combination made their app an instant success in many countries. They hit one million downloads in the first three months. This husband and wife team quickly realized that they needed to leave their current jobs and start their own company to support their massive fan base and growing number of advertising requests. They started Advanced Competition Simulations (ACS) on July 1, 2014 and hired their good friend, Samuel "Sam" Williams, to manage their growing databases.

### Current Datasets and Databases

When Rich and Jill started the company in 2014, they chose to use a PostgreSQL database for their customer and billing records since their app was already using SQLite and they were both familiar with PostgreSQL in their previous jobs. Over time, Sam setup a third-party advertising vendor database that contained all the data pertaining to the vendors that paid them for advertising space in their app. All their databases are SQL based, but some of their table names and aliases were

incorrectly labeled when Sam took over and chose a different naming convention. Their app database also contained unstructured data such as video clips of game replays.

In early 2015, Sam approached Rich and Jill and voiced his concerns about the stability of their servers and the quality of their data. Sam was also worried about the earthquakes that they had been having and suggested that they invest in new servers and an offsite location. Rich agreed and immediately began looking for the perfect location for their new data center. After a detailed risk assessment, Rich ultimately decided to move ACS's data center to Brentwood, Tennessee because, other than the occasional tornado, Tennessee was one of the safest places on his list. He also had a former college roommate, Levi Anderson, living there that he kept in touch with, and knew was looking for a career change. Rich hired Levi and had their new data center up and running by the end of 2015.

### Organizational Structure

After three years of continued success and the development of their second app, Rich and Jill decided to give themselves more appropriate titles. Rich assumed the role of Chief Executive Officer (CEO), Jill was appointed the Chief Information Officer (CIO), Sam was promoted to Information Systems Division (ISD) Director, and Levi was named the Director of Data Center Facilities Operations (DDCFO). ACS quickly grew from a four-employee company to more than 100 employees; to include one of the best development teams in the industry. ACS is now a \$250 million-dollar gaming company and generated \$125 million in revenue and \$40 million in profit last year. The majority of their revenue comes from in-app purchases, but a small portion comes from their advertisers. Rich has always viewed in-app advertisements as a nuisance, but a necessary evil in order to fund their expensive development team that is constantly working on new apps and improvements.

## 3. THE BIA PROJECT

On January 27, 2017, Rich returned from a conference where he saw a demonstration of Tableau, and decided that he needed to be able to visualize all the customer data that ACS had accumulated over the past three years to make better business decisions about the new app that the development team was working on.

Like Rich, many companies are realizing the potential of the data that they have been

collecting in their data centers. Companies have been using data to predict or “forecast” for decades, but recent updates to technology are allowing companies to speed up this process tremendously.

Software companies, like IBM and Tableau, are now focused on bringing business intelligence to users that typically do not query company databases for trends. They are achieving this by making the desktop software needed for creating custom queries and visualizations more mobile. Users can now use their iPad to answer key business questions on a moment’s notice.

Rich had already decided that he wanted to use the Tableau software, but his big decision now was whether he should hire a team to handle this internally or outsource this capability to a trusted vendor. Since Jill was still not happy that Rich decided to move forward with the project, she thought they should just outsource it to a company that does it every day so they could get back to the data center disaster recovery project. Rich believed that they should hire a team internally to have more control over their actions. At one point when Rich and Jill were discussing the options, Sam walked into their office and heard Jill make the comment:

"You know, you are always doing this Rich! You go off to these conferences and come back with these huge ideas and expect me to make them a reality! What makes you think we can handle hiring an entirely new team when the new app is taking up all of our resources and we don’t even have a disaster recovery plan? I am genuinely concerned about our data center."

After their heated debate and weeks of research, Rich decided that creating a business intelligence department within their company was the better decision in the long run.

Due to most of ACS’s resources being tied up on the new app project, Rich only gave Jill three months to implement Tableau, a project budget of \$250,000 and a recurring yearly salary budget of \$225,000 to hire a business intelligence and analytics team. Since Rich did not give Jill the budget to hire a well-established business intelligence expert, Jill hired an experienced business intelligence data scientist, Pradeep Dey, to head the department as the Director of Business Intelligence (DBI). Pradeep Dey had over 15 years of experience as a data analyst and 4 years of experience as a data scientist. As a data analyst, he was responsible for merging multiple datasets and answering questions that

the management team asked. Once he was promoted to data scientist, he was responsible for studying datasets independently and coming up with new questions that needed to be answered to push the company in a new direction.

Jill’s research suggested that the DBI should report directly to her instead of the Information Systems Division (ISD) Director. This would ensure that the business intelligence team was focusing on the direction of the company rather than just the IT department’s vision. Once Pradeep was on board and caught up to speed on ACS’s current setup, Jill and Pradeep hired two business intelligence analysts, Uche Ike and Michael Christopher, who report to Pradeep. Uche Ike recently moved to the United States from Nigeria, but had six years of experience as a business intelligence analyst. Michael Christopher was a recent college graduate with no experience, but was very eager to learn and was top of his class.

Rich and Jill’s longtime friend and current ISD Director, Sam, was currently the only employee that knew the intricacies of ACS’s databases and had experience with business intelligence software. He used Power BI in Microsoft Excel to connect to their existing desktop support ticket database to view IT employee metrics. This experience and relationship with Rich and Jill earned him the role as project manager from this point forward.

#### **4. PROJECT MANAGEMENT APPROACH**

Selecting the right project manager can make or break your project. In ACS’s case, Rich and Jill chose a trusted friend to lead the project instead of the newly hired Director of Business Intelligence. This did not bode well with the newly formed business intelligence team and caused the project to get off to a rough start. Pradeep, Uche and Michael all felt as if the executives did not trust them to do the job that they were hired to do. Pradeep voiced his concern to Jill in a casual hallway conversation, but Jill reassured Pradeep that Sam was chosen as project manager because of his vast knowledge of the current databases and that she would still rely heavily on Pradeep’s technical knowledge throughout the project.

Sam intended to approach this rollout in his usual agile way by keeping Rich and Jill in the loop every step of the way as the project progressed and altering the project based on their feedback. However, Pradeep favored a more traditional approach. He suggested that the project follow a

sequential series of steps that would include among other things: a requirements definition, planning, data cleansing and migration, testing, and deployment; as can be seen in Figure 1. Sam agreed to go along with this approach, because the project is internal, for an internal customer and as such may not be affected by incessant changes in scope. Sam may have agreed to Pradeep's iterative approach, but he was quick to let the project team members know his intention to inject this method with as much flexibility as possible.

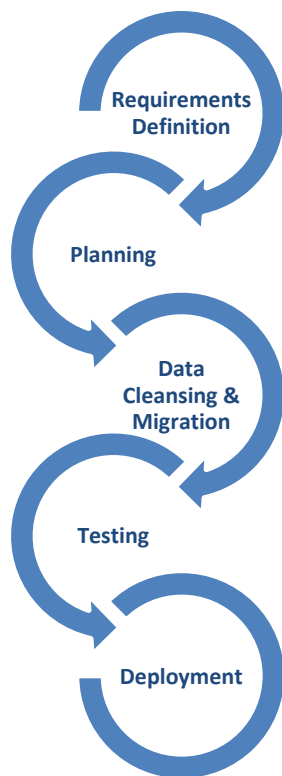


Figure 1: Pradeep Dey's Proposed Approach

### Requirements Definition and Planning

Sam still wanted to integrate as much agile project management into the project as he could, so he decided that daily scrums would take place in order for him to monitor the project more effectively. With this hybrid approach of iterative and agile, the goal was to breakdown the work into manageable chunks with limited alterations. After several weeks of brainstorming sessions, they agreed to phase the project. The team quickly defined the requirements, wrote the statement of work and planned the next phases.

The team emphasized effective communication as a necessary virtue for all the team members during most of the meetings and brain storming sessions. Due to the size of the team, they agreed

on making communication as open as possible. ACS runs a top-down communication structure, but the team figured that would cause unnecessary red tape that could hold up work significantly, so they decided to do away with structure as far as this project is concerned.

### Data Cleansing and Migration

Data cleansing is not a new idea to ACS. It was actually placed on the project priorities board a year ago by Jill and it had the third highest priority since the new app that the development team was working on would leverage only structured data.

Since ACS's data center contained structured, semi-structured and unstructured data, the first step in the data cleansing and migration phase was to improve the data quality by structuring the majority of the data. The objective was to weed out junk data that would generate wrong results and mislead the business. The team deemed it imperative that the validation and cleaning step scaled, since having a continuous data value chain required that the incoming data get cleaned immediately and at very high rates. This would require automating the process and possibly eliminating any form of human involvement with data entry.

Integration with either R programming or Python would help with procedural computation and data cleansing. R and Python programming are well equipped to handle this process. R helps to do the heavy statistical analytics. Their machine learning and artificial intelligence programs used the complex algorithms for detecting the repeatable patterns.

The second step in the data cleansing and migration phase was to architect a data model that the visualization would be based on. This would require a strong background in statistics and machine learning to build. The team was not interested in this skill as they intended to use R – statistics and a good number of the team members were very familiar with the language. On the other hand, their statistical background was not enough to build this model accurately. They needed someone who understood the business well enough to recognize whether the results of the mathematical models were meaningful and relevant. So they solicited more of Jill's involvement to cover the business knowledge area.

### Testing and Deployment

Data migration, testing and deployment of the Tableau software were intended to be a simultaneous process. Depending on how the test run turned out, they intend to go live with deployment immediately on their production server. However, the test server that was set up for the project was not configured correctly and the team had to move their migration efforts and testing onto the production server earlier than they had planned.

### 5. THE DATA CLEANSING SETBACK

The project appeared to be planned well and running smoothly until Lu Wei made his third status visit to ACS headquarters. Lu Wei is the Senior Tech Support Engineer with Tableau that is assigned to assist ACS with the implementation process. Lu makes weekly office visits to ACS's headquarters to check in on the project. On his third office visit Sam approached him in the hallway and asked how Tableau's side of the data cleansing was going. Lu looked at Sam with a surprised look and said, "We are not responsible for cleaning or formatting any of the data. I thought your team was handling that in the current phase?" Sam could not believe he was hearing this and immediately pulled Lu into his office, and had a lengthy discussion about the contract that was signed earlier in the project. Sam argued that he remembered asking Lu about it during the planning phase and Lu agreed that he would assist with a portion of the cleansing operation. Lu called Sam's bluff and pulled up the contract on his tablet. He specifically read the part that stated:

"ACS will be responsible for data cleansing and formatting all of their data correctly prior to the Tableau implementation and failure to do so could result in project delays."

Sam was not very happy, but the only choice he was left with was to have the team clean the data alone, which he thought would set them back several weeks. Luckily, the data cleansing operation did not take as long as expected, thanks to the expertise and productivity of Pradeep and the BIA team along with a decent amount of overtime spent working weekends. Once the cleansing was completed, Jill was beginning to think that this project would pay off after all. She was even able to check off the data cleansing project on the priorities board.

### 6. THE DISASTER

Jill is awoken at 4:30 a.m. again to the sound of her phone vibrating on her mahogany bedside table. She lets out a sigh and a moan before rolling over to see Levi's name on the caller ID. She says to herself, "if he asks me how to sync his OneNote again I am going to scream." Jill whispers, trying not to wake up her sleeping husband next to her, "Hey Levi, what's up?" Rich, is awakened by Jill's voice, and asks "is that Levi again?" "I just don't understand why he can't wait until our 7:00am video conference."

Levi responds, "Jill! I am sorry to wake you up but we have a big problem and Sam told me to call you. We've had a power failure and Koffi and I are having trouble getting the production server that was handling the migration to stay up for more than 10 minutes before it crashes again. I am not certain, but I think it is hardware related." Jill turns and looks at Rich and whispers, "do I have to say I told you so?"

Levi voiced his concerns over a year ago regarding not having a disaster recovery plan in place, and he recommended hiring an information security expert and setting up a hot site or outsourcing to an Infrastructure as a Service (IaaS) vendor. Jill agreed and added the "Disaster Recovery Project" to the project priorities board, but the project kept getting jumped by Rich's "great ideas," just like the BIA project. When Jill hung up the phone, she was quick to remind Rich about her comment when he decided to skip the other projects, "Didn't I say that we should at least finish the Disaster Recovery Project first?"

The entire team quickly set up a conference call to talk through the problem. Sam was the first to speak up with the following comment:

"The power failure could not have happened at a worse time, I mean, right during the migration phase, but I think we are going to be okay. Koffi is close to a solution."

Koffi Anan, the database administrator that Levi hired a few years ago, had been up all night researching the problem and was close to a solution but needed a little more time to confirm his theory. The team offered some advice to Koffi, but ultimately it was up to him to solve the problem.

## 7. CONCLUSION

As always, Rich prioritized his "great ideas" and fun projects over the boring IT projects. He assumed that the project would be relatively easy and cheap, but after several project setbacks and a disaster at their data center in Tennessee, he realized it was quite the opposite. This project ultimately put ACS's daily operations in jeopardy, and could put them out of business if they are unable to recover the failing production server.

## 8. DISCUSSION QUESTIONS

1. Would you have prioritized the projects differently?
2. Do you think Rich made the right decision to hire a business intelligence team instead of outsourcing the capability to a proven company?
3. Was the right person chosen to lead the project?
4. Did the Tableau integration go as planned or could it have been handled more effectively?
5. Could the disaster have been prevented or mitigated?

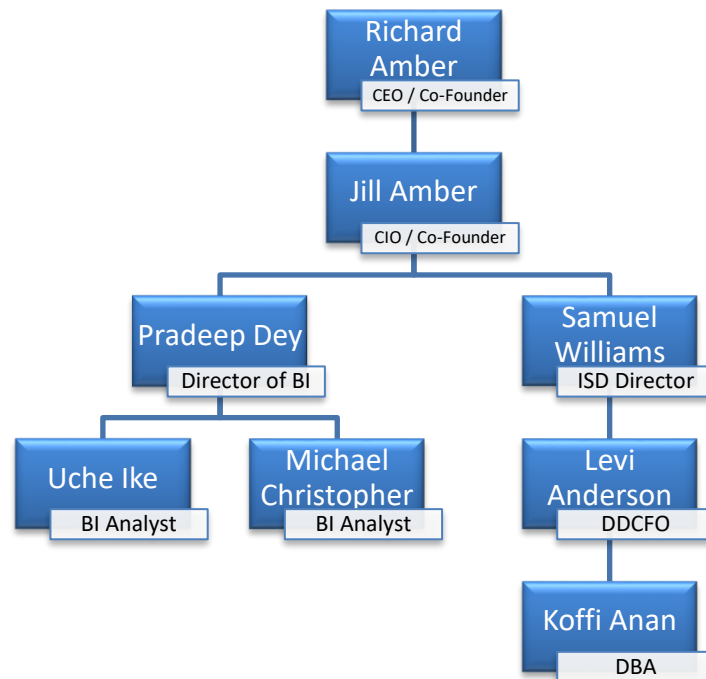
## 9. REFERENCES

- Ransbotham, S. (2015). Coca-colas unique challenge: Turning 250 datasets into one. *MIT Sloan Management Review*, 56(4), 0.
- Power, D. J. (2016). Data science: Supporting decision-making. *Journal of Decision Systems*, 25(4), 345-356.
- Englmeier, K., & Murtagh, F. (2017). Editorial: What can we expect from data scientists? *Journal of Theoretical and Applied Electronic Commerce Research*, 12(1), 5-I,II,III,IV,V.
- Elwer, P., & Corporation, I. (2008). Agile Project Development at Intel: A Scrum Odyssey, 1-14.
- Gwen Shapira. (n.d.). The Seven Key Steps of Data Analysis. Retrieved April 10, 2017, from <http://www.oracle.com/us/corporate/profit/big-ideas/052313-gshapira-1951392.html>
- Špundak, M. (2014). Mixed agile / traditional project management methodology – reality or illusion? *Procedia - Social and Behavioral Sciences*, 119, 939-948.
- Scavicchio, J. (2016, July 6). Tableau Pricing: How Much is Tableau and what are the Costs? Retrieved April 08, 2017
- Milligan, J. N., & Milligan, J. N. (2016). *Learning Tableau 10: learn how to create effective data visualizations with Tableau and unlock a smarter approach to business analytics: it might just transform your organization*. Birmingham: Packt.
- Acharya, S., & Chellappan, S. (2016). Integration of Tableau with R. *Pro Tableau*, 795-833. doi:10.1007/978-1-4842-2352-9\_11
- Weber, J. (2015). Collecting Data from Mobile Apps. *Practical Google Analytics and Google Tag Manager for Developers*, 221-230. doi:10.1007/978-1-4842-0265-4\_13
- Unstructured Data Analytics: Unstructured Data Analytics. (2015). *Win with Advanced Business Analytics*, 359-376. doi:10.1002/9781119205371.ch18
- Pham, T. (2016, March 15). The Company Who Made Clash of Clans is Worth More Than Evernote, Eventbrite, and BuzzFeed Combined. Retrieved from: <https://thehustle.co/the-company-who-made-clash-of-clans-is-worth-more-than-evernote-eventbrite-and-Buzzfeed-combined>



## Appendix

### ACS Organizational Chart



### Key Characters and Locations

#### ACS headquarters in Silicon Valley, California

Richard "Rich" Amber	Chief Executive Officer (CEO) and Co-Founder of ACS
Jill Amber	Chief Information Officer (CIO) and Co-Founder of ACS
Samuel "Sam" Williams	Information Systems Division (ISD) Director, formerly the Database Administrator
Pradeep Dey	Director of Business Intelligence (DBI)
Uche Ike	Business Intelligence Analyst 1
Michael Christopher	Business Intelligence Analyst 2

#### ACS Data Center in Brentwood, Tennessee

Levi Anderson	Director of Data Center Facilities Operations (DDCFO)
Koffi Anan	Current Database Administrator (DBA)

#### Tableau Vendor in Seattle, Washington

Lu Wei	Senior Tech Support Engineer assigned to assist ACS with the implementation of Tableau
--------	--

## Teaching Case

# Dragon Air: A Database Design Case

Donald Wynn, Jr.  
wynn@udayton.edu  
MIS, OM, and DS Department  
University of Dayton  
Dayton, OH 45469, USA

### Abstract

This case describes a real-world scenario based on a fictional new commuter airline seeking to develop a database to support their operations. Students are asked to design and develop a database to allow the airline to track flights, passengers, bookings, and more, based on a set of business rules, default data, and other requirements. The case has been designed for use in a database management course for students having little relational database experience other than the content of the course itself. Teaching notes containing suggestions and a possible solution are provided upon request.

**Keywords:** Database Management, Database Design, Entity-Relationship Diagrams, SQL.

### 1. COMPANY INFORMATION

Dragon Air is a new low-cost airline startup, with headquarters in Dayton Ohio. Initially, their focus will be on providing commuter flights between mid-size airports in the US. They have contracted your firm to develop a new reservation system to track flights, planes, crew, customers, and bookings. Your team's assignment is to develop the relational database to support the operational requirements of the new reservation system.

### 2. DATA REQUIREMENTS AND BUSINESS RULES

The company has identified a number of data elements and associated requirements, each of which needs to be included in the database. These rules concern the data associated with the flight destinations, flights, tickets, passengers, planes, and crew members. Additional information regarding the initial flight schedules, airport information, and their fleet of planes is included in the appendix.

#### Destinations

Dragon Air will initially fly to 11 cities in the United States. The cities are listed below. Eventually, it is expected that they will fly to many more cities

in the US and around the world. Each city is identified by a three-letter code assigned according to the IATA (International Air Transport Association) standard.

IATA Code	City
ABQ	Albuquerque NM
BUF	Buffalo NY
DAY	Dayton OH
GSP	Greenville SC
JAX	Jacksonville FL
MSY	New Orleans LA
PDX	Portland OR
SAC	Sacramento CA
SAT	San Antonio TX
TUS	Tucson AZ
TYS	Knoxville TN

**Table 1. List of Destinations**

#### Tickets and Bookings

All tickets are paid in advance, at the time of booking. Each booking has a unique 6-character booking ID, date, flight number, price, status

(booked, cancelled, abandoned, etc.), and any airport taxes paid. Also, each booking is subject to a 7.5% Domestic Passenger Tax (payable to the FAA) and a \$4.00 charge for each leg of the flight. Round trip tickets (i.e. in both directions between two cities) are counted separately. For instance, a round trip ticket between Buffalo and Dayton is recorded and billed as two flights: Buffalo to Dayton and Dayton to Buffalo.

Each booking is paid by a single customer, who may or may not be one of the flying passengers. Each booking is associated with a single payment, but a customer may pay for several bookings on the same payment. Dragon Air accepts payment by credit card or checks only. The relevant information for both payment types will need to be stored (e.g. card number, expiration, check number, account number, etc.).

At booking, the passengers may choose to select their seat in advance for \$25 extra. Regardless, when seats are assigned, this must be tracked in the database as well. Also, each passenger may prepay for their luggage. These costs are added to the booking price as needed. If the luggage is not prepaid, the passenger may pay at the airport. Baggage costs are shown below.

Item	Carry-on	Checked
Prepaid luggage	\$20	\$25
At the airport	\$50	\$50

**Table 2. Baggage Fee Chart**

### Passenger

### Information

We need to track the name, date of birth, passport country and number, email, mobile phone, and address information for each passenger. These passengers accumulate miles for each flight (regardless of who pays for it). Once they accumulate 50,000 miles, they become members of the Silver Flyers club. At 100,000 miles, they become members of the Golden Flyers club. This designation allows them to board flights earlier and to get upgrades to premium seats. Their flyer status is active for one year from the date it is activated (so this date needs to be tracked in their account). For each flight, we need to be sure that the passenger checked in at the airport on the scheduled date of the flight.

### Flights

For each flight between two cities, DA needs to track the flight number, departure and arrival cities, the miles for each flight, the days of the week on which they are scheduled, and the

scheduled departure and arrival times. They also need to track the airport taxes for both airports. Each flight occurs once on each scheduled day. For each individual flight, we need to track the actual departure and arrival times. We also need to know the plane which flew each individual flight, and the crew (pilot, co-pilot, and 4 flight attendants).

### Planes

For each plane, DA needs to track the FAA number, plane type (e.g. 737-700, 787, etc.), the current status (e.g. in service, out for repairs, retired, etc.) and the airport hangar where it is based. DA also wants to track the maintenance visits for each plane, including the date of service, the type of service performed, and the results of the service.

### Crew

DA employs six crew members for each flight: pilot and co-pilot and four flight attendants. The crew members on each flight need to be identified, including their advance schedule for the next few months. Although there are stringent regulations on how often a crew member can fly, this is not tracked or monitored within this database at this time. Currently, there are over 650 crew members in the database, including 180 pilots/co-pilots.

## 3. REQUIRED TASKS

The database must be able to support each of the tasks and outputs listed below.

1. Book a flight for a passenger.
2. Pay for a booking. Remember: multiple passengers may eventually be associated with a single payment.
3. Add a new 737-800 airplane (NX6471DA) to the fleet, with a capacity of 205 passengers.
4. Schedule crew members for a given flight on a given date.
5. Change a crew member on a given flight due to illness.
6. Change a passenger's flight to the next day.
7. Change the price of a flight.
8. Add a baggage fee for a customer at the airport.
9. Add new crew members (1 pilots and 1 flight attendant).
10. Change the departure date and time for a flight scheduled for a given date/time.

(For instance, assume there is a one hour weather delay.)

#### 4. Desired Outputs

Dragon Airlines' managers need to make decisions regarding their daily operations. To support their efforts, they have identified an initial set of information outputs that the database should support.

1. Scheduled flights, by date  
Show each flight scheduled on a given date. Show all flight information, including scheduled times and plane types.
2. Scheduled flights by date, restricted by airport.  
Show all flights scheduled on a given date, but only for a particular airport.
3. Capacity Report  
Show the capacity for all flights on a given date. This includes flight number, origin, destination, number of bookings, and capacity.
4. Passenger Manifesto  
Show all passengers scheduled and checked-in for a given flight.
5. Capacity Report Summary, by date range  
Show the total capacity and bookings for each date, summarized for a given start and end date.
6. Passenger history  
Show all flights taken by a given passenger, with miles and cost. Also, note if the passenger is a member of the Silver or Golden Flyers club, with the date their membership expires.
7. Bookings History  
Show all bookings paid for by a given customer. Note that this is different from the passengers that took a flight.
8. Silver Flyers / Golden Flyers  
Print a list of all Silver and Golden Flyers, sorted by the miles flown, including the passengers' name and total miles flown.
9. Booking Receipt  
Show all information associated with a given booking; i.e. booking customer, passengers, total price, total baggage fees paid in advance, scheduled times, and payment information.
10. Revenue Report by date range  
Show the total revenues booked on all flights between a given start and end date.
11. Revenue Report by flight  
Show the total revenues booked on a given flight.
12. Max Miles  
Show the top 5 passengers, in terms of the total number of miles accumulated on flights they have taken.
13. Max Flights  
Show the top 5 passengers, in terms of the total number of flights taken.
14. Passenger Schedule  
Show the flight information for all upcoming flights for a given passenger.
15. Plane Aging Report  
Show the total number of miles flown for each plane.
16. Crew Schedule  
Show the scheduled flights for a given crew member.
17. Crew Assignments  
Show all crew members scheduled for a given flight
18. Maintenance history  
Show all maintenance actions for a given plane.
19. Plane model history  
Show the number of maintenance events by the type of plane. If possible, distinguish routine maintenance from malfunctions, repairs, or other service issues.
20. Flight Discrepancy Report  
Show any flights where the actual departure/landing time differs from the scheduled times.

#### Editor's Note:

*This paper was selected for inclusion in the journal as an EDSIGCON 2017 Distinguished Paper. The acceptance rate is typically 7% for this category of paper based on blind reviews from six or more peers including three or more former best papers authors who did not submit a paper in 2017.*

## Appendix

### City Information

IATA Code	City	Time Zone	Latitude and Longitude
ABQ	Albuquerque	Mountain	35.0433° N, 106.6129° W
BUF	Buffalo	Eastern	42.9397° N, 78.7295° W
DAY	Dayton	Eastern	39.9025° N, 84.2218° W
GSP	Greenville	Eastern	34.8504° N, 82.3514° W
JAX	Jacksonville	Eastern	30.4941° N, 81.6879° W
MSY	New Orleans	Central	29.9922° N, 90.2590° W
PDX	Portland	Pacific	45.5898° N, 122.5951° W
SAC	Sacramento	Pacific	38.6951° N, 121.5901° W
SAT	San Antonio	Central	29.5312° N, 98.4683° W
TUS	Tucson	Mountain	32.1145° N, 110.9392° W
TYS	Knoxville	Eastern	35.8108° N, 83.9939° W

**Aircraft Fleet Information**

FAA Number	Plane Type	Capacity	Base
NX3377DA	737-800	160	ABQ
NX3521DA	737-700 ER	125	PDX
NX3594DA	767	200	DAY
NX3598DA	737-700	135	DAY
NX3609DA	737-700	135	SAT
NX3669DA	767	200	JAX
NX3704DA	737-700	135	ABQ
NX3897DA	737-700	135	SAC
NX3911DA	737-700 ER	125	TYS
NX4466DA	737-700	135	TYS
NX4592DA	737-800	160	PDX
NX4602DA	737-700	135	TYS
NX4770DA	767	200	DAY
NX4787DA	737-700	135	GSP
NX4845DA	767	200	JAX
NX4869DA	737-800	160	MSY
NX4928DA	737-700	135	BUF
NX4952DA	737-700	135	TYS
NX5052DA	737-700	135	GSP
NX5114DA	737-700	135	TUS
NX5165DA	737-700 ER	125	MSY
NX5641DA	737-700	135	JAX
NX5717DA	737-700	135	JAX
NX6070DA	737-700	135	SAC
NX6740DA	737-800	160	BUF
NX6784DA	737-700 ER	125	PDX
NX6813DA	737-800	160	SAT
NX6862DA	737-800	160	DAY
NX6864DA	737-700	135	DAY
NX6871DA	737-700 ER	125	TUS
NX7016DA	737-800	160	MSY
NX7037DA	737-700 ER	125	PDX

### Flight Schedule

Flight#	From	To	Departure	Arrival	Miles	Minutes	Price	Flight Days
484	Albuquerque	Buffalo	7:05:00 AM	12:32:00 PM	1585	207	\$ 310	MTWRF
167	Albuquerque	Dayton	11:33:00 AM	4:59:00 PM	1269	206	\$ 235	MTWRF
697	Albuquerque	San Antonio	2:00:00 PM	4:31:00 PM	609	91	\$ 148	
279	Albuquerque	Tucson	11:35:00 AM	1:05:00 PM	321	90	\$ 100	
515	Buffalo	Albuquerque	2:17:00 PM	3:44:00 PM	1585	207	\$ 279	MTWRF
257	Buffalo	Dayton	12:28:00 PM	1:51:00 PM	353	83	\$ 104	
183	Buffalo	Portland	4:02:00 PM	5:34:00 PM	2151	272	\$ 385	MTWRF
554	Buffalo	Sacramento	11:45:00 AM	1:14:00 PM	2236	269	\$ 384	MTWRF
180	Dayton	Albuquerque	8:22:00 AM	9:48:00 AM	1269	206	\$ 245	MTWRF
685	Dayton	Buffalo	9:20:00 AM	10:43:00 AM	353	83	\$ 107	
663	Dayton	Jacksonville	11:30:00 AM	1:00:00 PM	666	90	\$ 151	
133	Dayton	Sacramento	6:50:00 AM	8:18:00 AM	1985	268	\$ 373	
153	Greenville	Portland	5:27:00 PM	7:02:00 PM	2222	275	\$ 381	MWF
138	Greenville	New Orleans	5:24:00 PM	6:03:00 PM	570	99	\$ 136	MTWRF
506	Greenville	San Antonio	10:19:00 AM	11:50:00 AM	1010	151	\$ 193	MTWRF
204	Greenville	Knoxville	5:00:00 PM	5:29:00 PM	114	29	\$ 67	MTWRF
263	Jacksonville	Dayton	2:45:00 PM	4:15:00 PM	666	90	\$ 147	
537	Jacksonville	Portland	9:15:00 AM	11:54:00 AM	2424	339	\$ 406	MTRFSU
500	Jacksonville	Sacramento	5:51:00 PM	8:27:00 PM	2320	336	\$ 423	
158	Jacksonville	Knoxville	12:33:00 PM	2:06:00 PM	391	93	\$ 109	
462	New Orleans	Greenville	1:00:00 PM	3:39:00 PM	570	99	\$ 136	MTWRF
541	New Orleans	Portland	8:04:00 PM	10:48:00 PM	2047	284	\$ 345	MWF
579	New Orleans	Sacramento	7:20:00 AM	10:01:00 AM	1875	281	\$ 318	MTWRF
306	New Orleans	Tucson	1:15:00 PM	2:54:00 PM	1230	159	\$ 250	MTWRF
243	Portland	Buffalo	6:45:00 AM	2:17:00 PM	2151	272	\$ 406	MTWRF
190	Portland	Greenville	7:22:00 AM	2:57:00 PM	2222	275	\$ 389	MWF
108	Portland	Jacksonville	1:39:00 PM	10:18:00 PM	2424	339	\$ 448	MTRFSU

481	Portland	New Orleans	10:50:00 AM	5:34:00 PM	2047	284	\$	377	MWF
200	Sacramento	Buffalo	3:44:00 PM	11:13:00 PM	2236	269	\$	377	
455	Sacramento	Dayton	10:33:00 AM	6:01:00 PM	1985	268	\$	351	
583	Sacramento	Jacksonville	6:45:00 AM	3:21:00 PM	2320	336	\$	407	
608	Sacramento	New Orleans	2:20:00 PM	9:01:00 PM	1875	281	\$	344	MTWRF
270	San Antonio	Albuquerque	6:16:00 PM	6:47:00 PM	609	91	\$	137	
176	San Antonio	Greenville	2:20:00 PM	5:51:00 PM	1010	151	\$	196	MTWRF
587	San Antonio	Knoxville	2:20:00 PM	5:50:00 PM	946	150	\$	192	
555	San Antonio	Tucson	11:40:00 AM	1:11:00 PM	760	151	\$	168	
444	Tucson	Albuquerque	2:50:00 PM	4:20:00 PM	321	90	\$	102	
303	Tucson	New Orleans	4:39:00 PM	8:18:00 PM	1230	159	\$	230	MTWRF
414	Tucson	San Antonio	3:41:00 PM	7:12:00 PM	760	151	\$	169	
269	Tucson	Knoxville	2:04:00 PM	7:33:00 PM	1560	209	\$	279	
456	Knoxville	Greenville	8:00:00 AM	8:29:00 AM	114	29	\$	68	MTWRF
470	Knoxville	Jacksonville	9:15:00 AM	10:48:00 AM	391	93	\$	110	
598	Knoxville	San Antonio	11:05:00 AM	12:35:00 PM	946	150	\$	192	
187	Knoxville	Tucson	10:50:00 AM	12:19:00 PM	1560	209	\$	307	



## *Teaching Case*

# System Development and Data Modeling for Stevens' Wholesale Health Supplies

Dana Schwieger  
dschwieger@semo.edu  
Department of Accounting

Ziping Liu  
zliu@semo.edu  
Department of Computer Science

Southeast Missouri State University  
Cape Girardeau, MO 63701, USA

## **Abstract**

The case provides a realistic scenario that can be used in a systems analysis and design, accounting information systems or graduate level management information systems course. The case details the steps in the revenue and expense cycle processes of a health supply wholesale company providing students with the opportunity to map the business process data flows. Multiple assignment options are provided allowing instructors to select an assignment based upon course material coverage. Suggested assignments include the development of data flow diagrams, swimlane diagrams, a request for proposal and a response to the request for proposal, and database design and development artifacts. Instructor materials are available.

**Keywords:** Teaching Case, Process Design, Request for Proposal, Database Development

## **1. INTRODUCTION**

Josh Stevens, a recent Accounting graduate of Midwest University, stared out his office window on the ground floor of Stevens' Wholesale Health Supplies (SWHS). He had a big task ahead and he wasn't sure where to begin. He reflected back to when he was six years old and used to play in cardboard boxes outside that very window. Those were much simpler times and Stevens' Wholesale Health Supplies was a lot smaller then. He wondered, however, how much the company would have grown if they had not relied so much on pencil and paper.

His parents, Robert and Sarah Stevens, had known for some time that they needed to move more of their office processes to an electronic

format. However, finding the time to do all of the research necessary to purchase, or hire someone to build, a new system while running daily operations was nearly impossible for them to do. Josh knew the moment that he graduated from his parents' alma mater that his first project at the family business would be to start researching their inventory supply process in hopes of moving it from an Excel workbook to a more robust system.

## **2. BACKGROUND**

In the early 1990's, Robert Stevens was a health conscience marketing major taking classes at Midwest University. He was a regular at the University's gym where he met the love of his life, Sarah Smith, an exercise science and nutrition

major. After graduation, Robert took a job as a pharmaceutical representative and Sarah was employed by a local hospital as a registered dietitian. About a year later, Robert and Sarah got married and eventually started their own family with the addition of a precious baby boy they named Joshua, or "Josh" for short.

Robert had the gift of gab and was very good at selling pharmaceuticals. He was rewarded by a widening sales territory and growing customer list. With Robert's added responsibilities came a generous paycheck as well as more days away from home for sales calls. Sarah quit her job at the hospital to be a stay at home mom. Life seemed close to perfect except for Robert's phone calls home when Sarah would report on Josh's latest development milestone. Those things broke his heart. He was missing so much of Josh's life that they decided he should quit his traveling job and find something closer to home.

Robert hated the thought of losing the friendships he had developed with his longtime customers and the independence that he enjoyed as a sales representative. He also wasn't very fond of the spare tire he had developed around his midsection as a result of long hours in the car eating fast food. He was, however, looking forward to being with his family, joining a local gym and eating his wife's healthy home cooking.

As Robert started down the path for finding a new career and, now that he had the time, a healthier lifestyle, he realized that losing weight and gaining muscle were not as easy as they once were for him. He and Sarah also started running low on some of the health-related products he used to pick up for Sarah when he was in bigger cities on sales calls. These realizations made Robert start thinking if there might be a connection between the two and his future career.

### **3. STEVENS' WHOLESALE HEALTH SUPPLIES**

Robert discussed his thoughts with Sarah. With Robert's connections and Sarah's interest in diet and health, they came up with a plan for bringing vitamin supplements and health supplies to smaller markets. Robert busied himself working with the local small business development center to draw up a business plan. They quickly laid the foundation for their new company and applied for a small business loan. Stevens' Wholesale Health Supplies (SWHS) was born. What started out as a two person operation in the spare bedroom of their home had, in twenty years, grown into a

multi-state business employing fifteen sales people, three warehouse staff, five office workers and a recent new hire, their son Josh. Josh had grown up with the business. He went from playing with the boxes that had contained inventory shipments to breaking down boxes and taking out trash during summer vacations to now preparing to help design and oversee a system to monitor and order those boxes containing products to be shipped to clients around the country.

Although Josh was familiar with the company's operations, he didn't really know how everything worked. From his business classes, Josh realized that he needed to have a better grasp of the business processes. He wasn't getting any closer to that knowledge by staring out the window and thinking about boxes, so he decided that the best place to get started was where those boxes came into the picture, the warehouse.

### **4. WAREHOUSE**

Josh shook hands with Jerry in front of the warehouse manager's office. "Come on in Josh and have a seat. How can I help you?" Josh explained his first big project for the company and how he thought he should become familiar with the inventory process.

Jerry leaned back in his chair to get comfortable. "Well Josh, let's start even before the boxes get to the warehouse. SWHS has a list of approved vendors with whom we do business. Each one sells us a different product or set of inventory at an agreed upon price. When we get low on particular inventory, we send a purchase request listing the items and the quantity that we need to Sue in the main office. She has the office right next to yours. Sue usually orders inventory on Tuesdays and Thursdays; however, we fill out the purchase request form every morning at the start of the day based upon which items have reached their reorder point so that we can keep track of the requests being made. Most of our customers have standing orders which we fill each week. We also get special orders in, but we keep enough extra inventory on hand that we usually do not have to ask Sue to place a special order for us."

"I guess I should know this," Josh started, "but to make sure I understand, who are your customers and vendors?"

"That's actually a great question Josh," Jerry smiled. "Our name can kind of throw you off. We started out providing supplies to doctors' offices, health food stores, and vitamin supplement

stores. Once the inventory management system is developed, I think your parents are going to start trying to sell their products to individual people online. Our vendors are product manufacturers such as "Vital Health" for vitamins and minerals, "Herb's Herbs" for herbal supplements, "Personal Health Products" for personal care products such as Vitamin C and Retinol creams to reduce the signs of aging, and "Warrior Foods" who makes muscle building supplements and Paleo diet foods."

"How do you place the orders?" Josh asked.

"Our purchase request is a word processed form that we complete, save, and email each day to Sue. If there is an unexpected order, we usually receive an email from someone on the warehouse floor, asking us to place a special order. If the order needs immediate attention, we complete the purchase request form on the spot. However, most of the time the requested item is just added to the next mornings' purchase request. If Sue has to send a special purchase order to a vendor for a rush job, she will, but that is pretty rare."

"How does Sue contact the vendors?" Josh wondered.

"Sue has a word processed purchase order form that she emails to each vendor for the products that we purchase from that vendor. Sue saves a copy of each purchase order on her computer so that she can keep track of all of the orders that she has made. If we have a special order, Sue will call the vendor to give them a heads up on the order."

"What happens when you get the inventory?"

"Mike, our loading dock worker, usually gets at least one delivery every day. The deliveries usually come through UPS who also handles the shipments to our customers. Once he has checked the purchase orders and the shipping documents, he will unload the pallets of inventory from the truck and verify that the inventory matches the shipping documents. If everything matches, the purchase order and shipping documents are then sent to Wendy in the main office. She handles the accounts payable duties and pays each of our vendors. If the shipment does not match the shipping documents, then Mike notifies Sue of the discrepancy, sends her the documents and she contacts the vendor to get the situation straightened out."

"How do they get everything 'straightened out'?" Josh asked.

"That's something you'll have to take up with Sue when you visit with her. She handles working with the vendors. However, once the inventory is approved, unpacked and stocked, it's available for us to ship to our customers. Every morning, we receive picking lists for each of our customers. Sue prints the picking lists at the end of the day for the next day's work and leaves them in our mailbox at the main office. The picking lists are used to pull the items off the shelf that are used to fulfill the customers' orders. Once the items are picked off the shelf, they are packed inside a box for shipping along with a packing slip and the box is sealed. The bill of lading is affixed to the outside of the box and placed in the package waiting area to be shipped. Mike arranges the boxes by customer and location and works with the shippers to get the trucks loaded at the end of the day."

"What about product returns from our customers?"

"Those come to Wendy in the main office. You can talk with her when you go up to talk to Sue. I hope I was able clear up the warehouse process for you. I'll call Sue and let her know that you are coming."

## 5. PURCHASING

As Josh walked towards the main office, he wondered what the company might do if its data was more readily accessible. Sue was on the phone with a vendor when he arrived; but she waved him to a seat by her desk. When Sue set the phone back in its cradle, she said, "So I hear you're learning about our inventory management process. What do you want to know?"

"Well..." Josh started, "I would like to know about your role in the process of buying and selling inventory. I heard about the process from the warehouse perspective, but I'm not sure what goes on in this office."

"I work with the vendors," Sue began. "In fact, that was a vendor on the phone when you came in. I am renegotiating our contract with that vendor to try to get a lower price. Our sales of their products have been growing steadily and I had hoped to get a larger discount. They're thinking about it, so I'll talk with them later. However, back to your question. "I order products from a selected group of vendors. On Mondays, Wednesdays and Fridays, I work with the salespeople to collect new orders as well as handle returns. We don't have a lot of issues with our inventory, neither from our vendors nor from

our customers, but occasionally an issue arises. When there is a discrepancy in the items we receive or goods are damaged between the vendors' warehouse and ours, I work with your mom in Accounts Payable and with the vendors to get our bill adjusted or the items returned. When a customer's order is incorrect, I work with Wendy in Accounts Receivable and the customer to correct the situation. We may reduce the customers' bill, offer a discount, or ship another product depending upon the problem. Speaking of Wendy, she just walked by."

## **6. ACCOUNTS RECEIVABLE**

"Wendy!" Sue exclaimed. "Could you come in here a moment? Josh wants to know about your role in the inventory management process."

"Hi Josh!" Wendy said. "I collect customer payments. A lot of our customers have blanket orders that we fill on a regular basis. We send the invoice electronically and they usually pay within the ten day 1% discount window. Occasionally we get a one and done customer, but most are repeats. Our long standing customers are usually prompt to pay their bills, but occasionally a payment doesn't make it through. In those situations, we contact the customer and they usually get the situation corrected immediately. If a customer returns an item, we usually try to replace the item with inventory on hand, but if the customer does not want a replacement, we will first suggest putting the funds towards a future purchase. We will also refund the customer for their purchase, but we push the other options first."

## **7. ACCOUNTS PAYABLE**

"How do you pay your vendors for their products?"

"Your mom usually handles Accounts Payable," Wendy started. Most of our payments are electronic and just require a confirmation for payment, but she always checks the purchase order, packing slip and invoice to make sure that we are paying for what we received. If there's a discrepancy, she works with Sue and the vendor to figure out exactly what we need to pay. She always tries to take advantage of payment discounts. "

"What happens when the company needs to return something to a vendor?" Josh asked.

"That sometimes happens, but not very often. When a product has to be returned, she usually

contacts the vendor and returns the product. She then processes a debit memo to record the transaction so that we don't overpay for inventory we do not have."

"How do you collect payments from customers and handle returns?"

"Most of our customers have standing orders and are in the habit of paying for their products on a set day to take advantage of payment discounts. Our salespeople contact each customer on a regular basis to provide a more personal touch and confirm the sales quantity before the order is filled. Sometimes the customers need to order more of our products and sometimes less, it often depends upon the time of the year. People tend to think about their health at the beginning of the year when they make New Year's resolutions or are planning for a summer swimsuit and less about their health as the Thanksgiving and Christmas holiday parties arrive.

If a customer returns a product and wants their account credited rather than replacement items sent, we process a credit memo for that customer and their account is credited for the amount of the returned item. Since our customers can pay within ten days for a discount or 30 days with no penalty, we normally do not have to provide a refund. However, once we get our new systems in place and start selling to retail customers, I have a feeling that we will need to develop a process for handling refunds. Speaking of new systems, how are you going to get started?" Wendy asked.

## **8. GETTING STARTED**

"I think I have a better idea now of the processes that will use the new system," Josh reflected as he finished writing his notes. "You've been a great help! The first thing that I will do is to create some diagrams to illustrate what I understand the processes to be. Once I have the processes and data flow diagrams designed, I think I'll starting thinking about the data that will need to be collected in order to run the queries and process the reports that will need to be generated. I'll use those fields to design an ER diagram to give to the developers. I think I will also create a small database prototype to give to the developers as well. That way, I can run the system by everyone who will have to use it to see if that is what they need before we invest a lot of money into having a professional system built that can interface with an online storefront."

"Would it be okay if I sent the diagrams to you to see if I am understanding the processes correctly?" Josh asked Wendy.

"I'll be glad to take a look at them, but I may not interpret the diagrams the same way you do. Why don't you write up an explanation to accompany each diagram and send that as well? I can take a look at them, make comments on them and forward them on to the other folks in the process. Let me give you some examples of the forms that we use in the buying and selling processes. That should help too," Wendy suggested. (Appendix)

"That's a great idea Wendy. I'll get started when I get back to my office." Josh said as he shook Wendy's hand and headed down the hall."

## 9. ASSIGNMENTS

Students should assume the role of either Josh or a systems analyst consultant hired to assist Josh. The systems analysts' roles and responsibilities will vary depending upon the course and assignment. Examples of possible assignments are provided. Possible answers for some of the assignments are provided in the teaching notes.

### Request for Proposal

**Courses:** Systems Analysis and Design, Process Modeling, graduate level MIS course

Assume that Josh is ready to move from the paper-based system to a computerized system. He is wanting to solicit vendor bids to purchase the technical infrastructure that will be used to support the new system.

1. From SWHS perspective, develop the functional and technical requirements that would be included in a request for proposal (RFP).
2. From a potential vendor's perspective, develop the vendor's response to the RFP for the technical requirements. Essentially, you are proposing the hardware, software, networking, installation, documentation and training that will be required to implement the infrastructure.

### Process Modeling

**Courses:** Systems Analysis and Design, Process Modeling, Accounting information systems, graduate level MIS course

Josh would like to draw the functional processes out on paper and send them to the various departments to verify that he understood their business processes correctly.

1. Create swimlane diagrams modeling each of the processes. Write short narratives to accompany your diagrams to verify and support your interpretation of the processes.
2. Create data flow diagrams (DFD) to model each of the processes. Write short narratives to accompany your diagrams to verify and support your interpretation of the processes.

### Systems Analysis Design and Database Development

**Courses:** Systems Analysis and Design, Database Development, graduate level MIS course

Assume that Josh would like to develop a prototype of the system to be developed. He wants to develop the:

1. System development diagrams
2. Data dictionary
3. Database forms for data entry
4. Database navigation switchboard
5. Database queries to generate purchase orders, inventory reports, and customer invoices.
6. Database reports including purchase orders, inventory reports, and customer invoices.

## APPENDICES

### Stevens' Wholesale Supply Purchase Requisition Example

#### Purchase Requisition

Stevens' Wholesale Health Supply  
Submit to : Inventory Office, Attn: Sue

PR #: 170901200

PR Date: 08/27/2017

Requesting Department: Wholesale Sales  
Office: Sales – Region 1  
Requisitioner Name: Thomas Adams  
Phone: 555-1111

	Item ID #	Description	Quantity	Unit Price	Amount
1	SWSVC250100	Vitamin C, 250 mg – 100 count (box of 50)	2	\$45	\$450
2	SWSVC500200	Vitamin C, 500 mg – 250 count (box of 50)	2	60	600
3	SWSVC21K500	Vitamin C, 1000 mg – 500 count (box of 50)	3	70	1400
4	SWSVD500500	Vitamin D, 500 mg – 500 count (box of 50)	2	50	100
5	SWSMMCH001	Muscle Milk Chocolate (gallon jugs)	5	\$40	200
---					
---					
---					
---					
	Total				\$4650

Date Needed: 09/05/17

Special Instructions:

Ordered by: \_\_\_\_\_ Authorized by: \_\_\_\_\_ Date: \_\_\_\_\_

### Stevens' Wholesale Supply Purchase Order Example

#### Stevens' Wholesale Health Supply Purchase Order

100 North Main Street  
Midwest Town, IL 62901  
(618) 555-1212

PO #: 170901200

Vendor Name:	Personal Health Products
Vendor Street:	Warehouse Street
Vendor City, State Zip	St. Louis, MO 63101

P.O. Date	Requisitioner	Shipped Via	F.O.B.	Terms
9/1/2017		UPS		1% in 10 Net 30

	SKU/Item ID #	Description	Quantity	Unit Price	Amount
1	PHPVC250100	Vitamin C, 250 mg – 100 count (box of 50)	10	\$45	\$450
2	PHPVC250200	Vitamin C, 250 mg – 250 count (box of 50)	10	60	600
3	PHPVC250500	Vitamin C, 250 mg – 500 count (box of 50)	20	70	1400
4	PHPVC500100	Vitamin C, 500 mg – 100 count (box of 50)	10	60	600
5	PHPVC500500	Vitamin C, 500 mg – 500 count (box of 50)	20	80	1600
6					
...					
...					
	Total				\$4650

Delivery address:	Delivery Date: 9/08/2017
Stevens' Wholesale Health Supply Warehouse 101 North Main Street Midwest Town, IL 62901	

Authorized by: \_\_\_\_\_

Date: \_\_\_\_\_

**Stevens' Wholesale Health Supply Invoice Example**

**9/1/2017**

**Invoice No. 1234**

To  
Vitamin See Optical Shop  
229 Broadway  
St. Louis, MO 63101

Ship To  
Vitamin See Optical Shop  
229 Broadway  
Springfield, MO 63101

Salesperson	P.O. Number	Requisitioner	Shipped Via	F.O.B. Point	Terms
Tom Adams	VSO201708200	Dr. Isaacs	UPS	Destination, Freight prepaid	1% in 10 Net 30

	Quantity	Description	Unit Price	Total
1	50	Healthy Vision Vitamin Supplement	5.00	\$250
2	25	Lutein	2.00	50
3	50	Vision Essentials Vitamin Supplement	7.00	350
4	25	Vitamin C, 250 mg - 500 count	2.00	50
...				
...				
		Subtotal		\$700
		Sales Tax		
		Shipping & Handling		45
		Total Due		\$745

Total due in 30 days from invoice date or within 10 days for 1% discount.

Thank you for your order!

We accept checks, cash and all major credit cards.

For questions about your order, please contact Wendy at (618) 555-1112.

**Stevens' Wholesale Supply**

Tel (618) 555-1111      100 North Main Street  
Fax (618) 555-1112      Midwest Town, IL 61901



### Stevens' Wholesale Health Supply Inventory Report Example

#### Inventory Report

Stevens' Wholesale Health Supply

Date: 8/31/2017

	Item ID #	Description	Quantity	Unit Price	Amount
1	SWSVC250100	Vitamin C, 250 mg – 100 count (box of 50)	20	\$45	\$900
2	SWSVC500200	Vitamin C, 500 mg – 250 count (box of 50)	18	60	1080
3	SWSVC21K500	Vitamin C, 1000 mg – 500 count (box of 50)	8	70	560
4	SWSVD500500	Vitamin D, 500 mg – 500 count (box of 50)	19	50	950
5	SWSMMCH001	Muscle Milk Chocolate (gallon jugs)	15	40	600
6	SWSHVV00001	Healthy Vision Vitamin Supplement (box of 50)	12	60	720
7	SWSL0000001	Lutein (box of 50)	14	50	700
8	SWSVEVS0001	Vision Essentials Vitamin Supplement(box of 50)	15	70	1050

## *Teaching Case*

# Formula One – a database project from start to finish

Anthony Serapiglia  
Anthony.Serapiglia@stvincent.edu  
CIS Department  
St. Vincent College  
Latrobe, PA 15650

### **Abstract**

The following is an applied database scenario based on a single season in the FIA Formula One (F1) World Championship of auto racing. This scenario builds database understanding and skills through data modeling, data acquisition, creation of a database schema through a database management system, query construction, and report creation. In the United States, Formula One falls squarely in the realm of “niche” sports, yet is regarded as one of the largest and most popular sports world wide. Various agencies have measured television viewership of F1 to be the largest for a seasonal sporting league in the world and only surpassed by quadrennial events such as the World Cup of Soccer and the Olympic Games. This dynamic allows for a great opportunity for undergraduate students. While most students will not be familiar with F1, they will have a basic understanding of racing in general. Also, while not in the consciousness of most Americans, the amount of sources of data for F1 abound worldwide. This creates a fertile environment abundant of data and the opportunity for easy entry to understanding of the environment by the novice. Utilizing this scenario will require that students learn about the world that the data is describing, enriching the experience with the personalities and energy of a world class sporting environment. The environment also provides multiple points of focus for modularization, but taken as a whole, allows for a full database creation experience, from “start” to “finish” - from modeling to reporting with all aspects in between.

**Keywords:** Data modeling, SQL, ER Diagram, Relational Database, data extraction, data manipulation

### **1. INTRODUCTION**

One of the primary concerns when beginning database work is for students to appreciate the value of building an understanding of the subject matter the database is to be built around. There are many definitions for a database, but many hold the following in common - a database is a collection of inter-related data that describes objects or events that exist or occur in a defined world. One essential that is not often stated and assumed to be accepted, but is often neglected and forgotten is that a database must have a purpose. There must be some reason that it has come into being. Without a clearly defined

purpose, those connections that define the interrelations of the collected data become weak, or nonexistent, allowing the structure to deteriorate and result in inefficient or unusable.

The purpose of this scenario is to build a database encompassing data from the current Formula One racing season. A user will be able to query this database to answer several basic questions of the data. Questions such as, “Who has won the most races?” or “What are the current team standings?” or “Who has qualified in the top ten the most times during this season?”

To create such a database, several steps are involved as evolutionary steps. First, an understanding of the world of F1 must be obtained so as to be able to identify what data is necessary, what is just useful, and what is not needed at all. To do this, sources for data and background context must be found. Second, a model needs developed to identify the entities that populate this world and their descriptors or attributes. Third, a schema needs to be developed within a database management system that will hold the F1 data. Fourth, a methodology must be developed to acquire data from sources outside of the database, that data transformed to fit the requirements of the database schema, and then loaded into the database. Fifth, queries will be necessary to select desired data from the database. Finally, sixth, reports can be developed to present the results of queries in a readily digestible and professional manner.

## 2. THE WORLD OF F1

Formula One (F1) is recognized by the Fédération Internationale de l'Automobile (FIA) as the highest level of open wheeled single seat auto racing in the world. Grand Prix racing can trace its roots to early road races in France, the earliest of which that utilized the title "Grand Prix" is recognized to be a 1901 event at the *Circuit du Sud-Ouest* around the streets of Pau (Rushby, 2011). Modern F1 history begins in 1950 with the first official World Championship for Drivers. Originally the championship season consisted of six events in Europe and the Indianapolis 500. The European events all had grown out of national championship races that had begun to appear prior to World War II. The first season saw Grand Prix races in Great Britain, Monaco, Switzerland, Belgium, France, and Italy. Other races were held as exhibitions, but did not count toward the championship. Non-points earning races would continue to be seen for many years until 1993. The number of races in the championship has varied from year to year, with the highest number being 21 events in 2016. While Europe remains the ancestral home of the series, races are now found across the globe including Asia, South America, North America, and the Far East. The 2017 season began in Australia in March and will end in Abu Dhabi in November.

Each season, two primary championships are contested, the World Constructors' Championship and the World Drivers' Championship. For the 2017 season, 10 teams will compete for the World Constructors' Championship. One team, Ferrari, has participated in every F1 season since 1950. The newest, Haas F1, enters its second season

and was the first US based F1 team in over 20 years. Each team consists of two cars and two drivers. The drivers compete for the World Drivers' Championship. Both championships are amongst the most prestigious trophies in the sporting world and there are many times that the Constructors trophy has gone to a team other than the team for the championship winning driver.

Round	Grand Prix	Circuit	Date
1	Australian Grand Prix	 Melbourne Grand Prix Circuit, Melbourne	26 March
2	Chinese Grand Prix	 Shanghai International Circuit, Shanghai	9 April
3	Bahrain Grand Prix	 Bahrain International Circuit, Sakhir	16 April
4	Russian Grand Prix	 Sochi Autodrom, Sochi	30 April
5	Spanish Grand Prix	 Circuit de Barcelona-Catalunya, Barcelona	14 May
6	Monaco Grand Prix	 Circuit de Monaco, Monte Carlo	28 May
7	Canadian Grand Prix	 Circuit Gilles Villeneuve, Montreal	11 June
8	Azerbaijan Grand Prix	 Baku City Circuit, Baku	25 June
9	Austrian Grand Prix	 Red Bull Ring, Spielberg	9 July
10	British Grand Prix	 Silverstone Circuit, Silverstone	16 July
11	Hungarian Grand Prix	 Hungaroring, Budapest	30 July
12	Belgian Grand Prix	 Circuit de Spa-Francorchamps, Stavelot	27 August
13	Italian Grand Prix	 Autodromo Nazionale Monza, Monza	3 September
14	Singapore Grand Prix	 Marina Bay Street Circuit, Singapore	17 September
15	Malaysian Grand Prix	 Sepang International Circuit, Kuala Lumpur	1 October
16	Japanese Grand Prix	 Suzuka International Racing Course, Suzuka	8 October
17	United States Grand Prix	 Circuit of the Americas, Austin, Texas	22 October
18	Mexican Grand Prix	 Autódromo Hermanos Rodríguez, Mexico City	29 October
19	Brazilian Grand Prix	 Autódromo José Carlos Pace, São Paulo	12 November
20	Abu Dhabi Grand Prix	 Yas Marina Circuit, Abu Dhabi	26 November

Figure 1 – the 2017 Formula One Race Calendar.

Points are awarded at each race. For many years, 1960 through 2002, F1 was one of the stingiest points allotments of any sport. Only six places earned points. This often led to dominant teams quickly creating insurmountable leads and settling the championship well before the end of the season. In an effort to combat this situation, more points were allotted to more finishers with less gap between. The current point system was put in place for the 2010 season and allows for 10 finishers to earn point on a scale of 25 for the winner, 18 – 2nd, 15 – 3rd, 12 – 4th, 10 – 5th, 8 – 6th, 6 – 7th, 4 – 8th, 2 – 9th, and 1 for tenth.

A race weekend involves three days. Each event will allow for two “free” practice session on Fridays. (The exception being Monaco which begins on Thursday to allow for the traditional Friday market day in the principality...). Saturday begins with a third “free” practice followed by qualifying. Sunday is race day. “Free” practices earned the name due to the fact that although timing and scoring will display results of the sessions, they do not count towards anything. The teams are “freed” of the normal race regulations and allowed to test new parts, as well as occasionally test new drivers. It is not uncommon to see a Friday test driver take part the first session before giving way to the normal race driver for the rest of the weekend.

Qualifying formats in F1 have also evolved over the years in an effort to both reduce costs and to produce a better show for fans. The current system of Qualifying consists of three separate sessions, Q1, Q2, and Q3. All cars have an opportunity to participate in Q1. At the end of 18 minutes, the top 15 move onto Q2. After a short break, Q2 lasts 15 minutes where only the fastest 10 drivers move onto Q3. The fastest driver in Q3 earns the pole position, or the right to start from the first position in Row 1 at the start of the race.

Each race has a different lap length and overall length of the race. Most races are around 300KM and last approximately 90 minutes. At the end of each race, the top three drivers make up the podium ceremony. Embracing a rich nationalistic history, during the podium ceremony the national anthem for the winning driver is played followed by the national anthem for the winning car constructor. In 2017, the 10 teams represent 8 different countries of home base, while amongst the 20 drivers 15 nations are represented.

**Section 2 Assignment:** Write a 1 to 2 page summary explaining the world of Formula 1. Include: What defines a season. Who are the major participants (Teams, Drivers)? How does a

race weekend run? What are the two major championships being fought for? How are results of practice session, qualifying, races reported? Include at least three websites as sources for data for results as well as background information about F1.

### 3. DATA MODELING

Following the work done in section one, a basic understanding of the actors/dynamics of the world of F1 should have been achieved. From this background, a model of what data describes this world can be determined.

Before starting in to build out a database in the database management system (DBMS), it is essential to organize and plan ahead of time. A data model helps to visualize the data and to aid in identifying the relationships that exist between the different groups. Data models can be very simple, or more complex depending on the needs of the developer, the data involved, and the eventual end “product” that is to be strived for.

Entities should be identified, the general categories that will contain multiple entries and their descriptors or attributes. In the case of F1, the race calendar can be a starting point. “Events” or “Races” can be used as a label for this entity. A listing can then be developed that includes the pieces of data that describe each event. What number the event is on the calendar, what country the race is located in, the dates the event takes place, the track name, the city name, the “official” sponsored title of the event, etc... can all be included as attributes.

Following this identification, Entity Relationship (ER) Diagrams should be developed to visually display the connections between these collections of data. Does a driver appear in a collection of drivers? Teams? Race Results? Practice Results? Qualifying results?

#### Section 3 Assignment:

Begin to identify data points in the world of F1. In step 1, create a document that identifies entities that exist. Under each entity, bullet out the attributes that serve as descriptors for each item that would belong in that collection. Begin to include a data type to describe what form the data will be in for each of the attributes. In step 2, develop a preliminary ER diagram of your entities. Identify attributes that the entities have in common.

#### 4. BUILDING OUT THE DATABASE

Different database management systems can refer to the software structures that make up the tables and other objects within the database in different ways. For this scenario, MySQL will be used as a reference for general vocabulary describing the structures within the DBMS. Thus "schema" will refer to the collection of tables, saved queries, and other objects that function as the "F1 Database" for this scenario.

Most DBMS packages include a visual design component. It is important to note that while these GUI (Graphical User Interface) interfaces can be extremely convenient, in almost all cases there will be certain advanced tasks that will still need to be completed through issuing standard SQL commands. It is always a good idea to become familiar with the specific SQL commands to issue that will satisfy even the most fundamental tasks such as table creation.

A common best practice is to save SQL commands in a plain text file for later reference. It is always a good idea to keep a running log of each command committed to the database in the case that it may need recreation from scratch. Having these commands saved off will allow anyone to be able to retrace footsteps and ensure that no steps are missed in the worst case recovery scenario.

This is also the time for verifying several hard choices that are made in the data modeling stage. When transitioning from "paper" to "practice" the absolute need to adhere to several database rules comes into play. The F1 world allows for several options when choosing primary keys for different tables. It also presents several "red herrings" that can be tempting to use, but could lead to duplication in several instances. Can the car number be used as a primary key within the Drivers table?

##### **Section 4 Assignment:**

Build out a shell of a database structure within a DBMS that fulfills your data model. For each step, save the SQL commands that were utilized in creation of tables and fields, the setting of datatypes, and identification of primary keys. Important constraints should also be identified, such as no-nulls. These SQL commands should be collected on a text file and clearly identified with comments defining what each section of commands is for.

#### 5. ETL

With a shell of a database built, it must be filled with data. Just like many other sports, F1 has a wealth of available statistics available over the Internet. One of the best sources is the official F1 web site - <https://www.formula1.com/>. In the results section of the website, a report is available with the summary statistics for each of the five sessions that cars are on track.

One of the most common and most important tasks of database work is managing an ETL process – Extract – Transform – Load.

In the case of a basic F1 database structure, a handful of common tables will prove to be essentially static, that is once data is loaded there will be little change or updates during the season. A table for events, the table containing teams, and the table for drivers are all relatively static. They are all also relatively small. For each it is manageable that the data they contain could very well be directly keyed in. While this is possible, it is not good practice. Anyone who takes this route may well be tempted to continue in this fashion. They will soon find it will be much to their advantage to develop a system to prepare datasets of session results and to import them through SQL scripts. While a team table may have only ten records, through the course of a season, twenty drivers per session with five sessions per event and twenty events per season will result in at least 2,000 records of results. While this may not be considered "Big Data" by some, it does demand that a more efficient method than direct keying of data be found.

In deciding a method of extraction and transformation there is no singular "best" way, although there are some "better" ways. Much of this depends on the source and shape of the data source that is providing the raw material. For most, an approachable method can consist of a copy and paste of data from the <http://formulaone.com> website into a spreadsheet. The spreadsheet can be arranged to transform the data, performing replacement of event, driver, and team names with corresponding foreign key identifiers for example. These transformed data sets can then be saved off as comma delimited (CSV) files which in turn can be imported into corresponding tables within the database. More advanced methods of ETL can include data scraping utilities, or custom programs in Java or C++ to collect data from a target and produce an arranged CSV file. Loading of data can also be broken into several steps. It is not always a good idea to load data straight into the final destination table. Holding tables can be

created as an initial import destination where data can then be validated and further moved to its final resting location.

Depending on the time of year, several races will have already been completed, with several more to take place. The ETL procedure can thus be tested with exiting data, and the process validated with new data as it arrives with the results of new races.

#### **Section 4 Assignment:**

Locate a good source of data that includes all necessary items to fill your database. Develop a method of capturing data, transforming it into a format that can then be loaded into your database. Detail a set of instructions for your procedure. The instructions must be specific and clear enough that another person unfamiliar with your database and the data will be able to follow them to a positive result. Load all possible data into your database.

### **6. QUERY THE DATA**

Data has potential. If allowed to simply sit in its static state within a database, that potential value will never be realized. To unleash the value of the data, it must be processed into information. Processing of data comes in many forms, from simply separating it from other pieces of data to advanced analytics and mathematical formulas. The purpose of the data processing is to answer a question – information informs. Data has potential, information is kinetic. Ask questions of your data.

The relational data model separates several key components that are normally found together. A normalized F1 database design would not have a table that contained a record that included a team name with the first and last names of their two drivers for example. Once data is loaded into the database, verify and validate the structure of the data model by performing several basic queries.

Simple queries to start with will look to combine data in just two tables. The example of a query providing a list of all teams with their drivers is a good example. To perform this query properly a basic SQL JOIN statement can be used to link the two tables together.

Once the basic structure has been confirmed for two table combinations, further test structure by producing a query that ties together four tables to display a race or practice session result. This should require a join between tables for events, teams, drivers, and session or race result table.

#### **Section 6 Assignment:**

Produce SQL queries that will pull data to answer the following questions: Who are the drivers for each team? What were the results from the second practice session of the fifth race of the season? What is the current standings for the Constructor's Championship? How many drivers have started on pole position and how many times for each? What races have been run and who were the winners to this point in the season? Who are the top ten drivers in terms of laps completed during practice sessions?

### **7. OPTIONAL - REPORTING**

(This section optional dependent on access to reporting software)

One of the biggest differences between enterprise level DBMS packages and a personal database program is the inclusion of a reporting component. Users of programs such as Microsoft Access will find relatively robust and user friendly modules for the production of reports generated from the data within the database. Traditionally, enterprise level database systems such as Oracle, Microsoft SQL (MSSQL), and MySQL have not included such features. Third party applications such as Crystal Reports and open sourced coding through PHP web pages have been utilized dependent on the need and the end audience.

In more recent releases of MSSQL, Microsoft has developed a report server that includes a free report builder component that can be used independently of MSSQL and the server component. A leading consumer application that is a competitor to Crystal Reports is DBExtra.

Reporting out of a database is a critical component to realizing the ultimate potential value of the data it contains. Whether the data is pushed out to recipients on a scheduled basis, pulled on demand by power users, or distilled for easy digestion through dashboards – reporting is often the final step in turning data into information by making it consumable by decision makers.

#### **Section 7 assignment:**

Based in the queries completed in the previous section, produce reports that will display results of each practice session, qualifying session, and race. Follow with preset reports that will display current driver and constructor standings.

## 8. OPTIONAL – FINAL ASSIGNMENT

With the recent takeover of the F1 commercial rights by Liberty media have opened up a opportunity for a special 21<sup>st</sup> race to be added to the 2017 season! Held in (insert exotic location somewhere in the world here) the race went off in secret until now. F1snoops.local (a local webserver to your location) has received race reports for the "lost weekend".

Based on all of the experience gathered through the past several assignments:

- Import the results of the event weekend into your F1 Database.
- Update any new drivers participating in Friday Free practice 1 in the Drivers table
- Produce reports for each of the five sessions of the weekend
- Produce updated reports with the final driver and constructors championship standings
- Provide a copy of the procedure followed in your ETL process.

## 9. REFERENCES

- A brief history of Formula One. (n.d.). Retrieved May 3, 2017, from <http://en.espn.co.uk/f1/motorsport/story/3831.html>
- Elmasri, R., & Navathe, S. (2011). Fundamentals of database systems. Boston: Addison-Wesley.
- Official Formula One Website. (n.d.). Retrieved May 12, 2017, from <https://www.formula1.com/>
- Rushby, K. (2011, April 29). Speeded date: motor racing returns to Pau. Retrieved May 10, 2017, from <https://www.theguardian.com/travel/2011/apr/30/pau-france-race-circuit>
- Spurgeon, B. (2016, March 18). Is Formula One Still on Top? Retrieved April 21, 2017, from <https://www.nytimes.com/2016/03/19/sports/autoracing/is-formula-one-still-on-top.html>
- Williamson, Martin "A Brief History of formula One" <http://en.espn.co.uk/f1/motorsport/story/3831.html>