BUSINESS INTELLIGENCE FINAL PROJECT WORKSHOP

Overview

During today's class, you will be given time to work on a new dashboard that will serve as this course's final project. You may choose to use Power BI or Tableau. The first stage of this project is to build your dashboard. In the next class, you will have an opportunity to reflect on the first draft of your dashboard and make any necessary changes. The last day of the Business Intelligence course will consist of presenting your dashboard to the class. You will be graded both on the content of your dashboard and your presentation.

We recommend you review the dashboard's requirements and grading rubric (found on the last page of this document) before selecting your data set.

Task 1: Selecting a Data Set

The first task of the project is to select a data set for your dashboard. We have provided six different data sets from which to choose. Each data set is given a rating from easy to challenging. There is no reward for selecting a more difficult data set. You are responsible for your own learning and can choose to challenge yourself as much (or as little) as you like.

The description of the dashboard goal is intentionally vague. This gives you more flexibility in how you choose to create the dashboard and what data you decide to display.

Data Set 1: Sports Difficulty Rating

Difficulty Rating: Easy
Goal: Create a dashboard that allows users to explore the skills required to excel at the most difficult sports.
Description: This data set includes 60 sports ranked by ESPN based on their level of difficulty and required skills. The data set contains 13 columns and 60 rows. The data set can be found here, while the original ESPN article (with descriptions of each column) can be found here.
Guiding Questions:

- What is the most difficult sport?
- Which sports require the most strength, speed, etc.?
- How do two (or more) sports compare in required skills?

Data Set 2: Marvel vs. DC Movie Ratings

Difficulty: Easy/medium
Goal: Create a dashboard that allows users to compare Marvel vs. DC superhero films.
Description: This data set includes information about every live-action Marvel or DC film released between 1944–2020. The data set contains 19 columns and nearly 90 rows. The data set and descriptions of the column names can be found here.
Guiding Questions:

- Which brand (Marvel vs. DC) has the higher average IMDb rating?
- Which brand has been more profitable over the years?
- How has the average rating of each brand changed over time?

Data Set 3: NASA Astronauts 1959–2013

Difficulty: Medium
Goal: Create a dashboard that displays information about the people chosen to be NASA astronauts between 1959–2013.
Description: This data set includes information about every NASA astronaut between the years 1959–2013. The data set contains 19 columns and nearly 350 rows. The data set can be found here.
Guiding Questions:

- Which universities have produced the greatest number of astronauts?
- What is the most common undergraduate major for astronauts?
- How has the gender makeup of astronauts changed over time?

Data Set 4: Oscar Best Picture Winners and Nominees

Difficulty: Medium
Goal: Create a dashboard to display information about the films that have been nominated for or won the Academy Award for Best Picture.
Description: This data set includes information about every film that has been nominated for or won the Academy Award for Best Picture. The data set contains 30 columns and nearly 500 rows. The data set and descriptions of the column names can be found here.
Guiding Questions:

- Which film studio has won the greatest number of Oscars?
- What is the average runtime for winning pictures compared to non-winning pictures?
- Which movie genre has won the Academy Award for Best Picture most often?

Data Set 5: Earthquakes 1965–2016

Difficulty: Medium/challenging

Goal: Create a dashboard that displays information about global earthquakes between 1965–2016.
Description: This data set includes information collected from the National Earthquake Information Center (NEIC). It has compiled a list of every major earthquake reported worldwide with a magnitude of 5.5 or greater between 1965–2016. The data set contains 21 columns and more than 23,000 lines of data. The data set can be found here.
Guiding Questions:

- How has the intensity of earthquakes changed over time?
- Has there been an increase or decrease in the average number of earthquakes per year?
- Is there a certain time of year when earthquakes are more likely to occur?

Data Set 6: US Census 2020

Difficulty: Challenging
Goal: Create a dashboard that displays demographic information of a state or territory, or create a dashboard comparing two different states or territories.
Description: This data set includes information collected during the 2020 United States Census. It is *strongly recommended* that you select only one or two states or territories for this project due to the full data set being exceptionally large. Working with this data set will be a challenge, but many interesting things can be done with it. The data set contains nearly 400 columns and thousands of rows of code (varying by state or territory). The full data set and descriptions of each column name can be found here. CSV files for individual states and territories are near the bottom of the page.
Guiding Questions:

- What is the racial makeup of the state or territory?
- What portion of the state's/territory's population is currently incarcerated?
- How do the demographics of one state or territory compare to another?

Task 2: Data Exploration

Now that you have selected your data set, let's spend time exploring it. Download the data set and open it in a program like Excel. You've done data exploration before, but we'll still give you a few things to think about as you explore.

Questions to consider during exploration include:

- Will any data cleaning be necessary before uploading the data to Tableau/Power BI?
- Are any of the columns superfluous?
- For whom was this data set created? What is its intended use?

Record any observations or notes about the data set below:

There are a lot of missing values in the data. I will need to clean the data by adding missing values and correcting any wrong information. There are many columns that are unnecessary. I ended up deleting all the rating count columns, the upright/spilled column, the description of the movie, the actors and directors information to name a few. I did this because I want to focus on the other information so that my dashboard doesn't get too cluttered.

Task 3: Planning

Planning can be tedious, but it is necessary and will make building your dashboard easier. Let's revisit everyone's favorite guiding questions: the 5Ws and 1H. Complete the chart below with information about your future dashboard.

Who?	This dashboard is for people who want to know more about the nominees and winners of the "Best Picture" award.
What?	The data set shows the main features of a film and then adds ratings. My dashboard will show the film title, oscar year, and average rating.
Where?	The data comes from the Academy Awards, IMDB, and Rotten tomatoes.
Why?	The academy award for "Best Picture" is the highest honor a film can receive.
When?	The data set takes place from 1932 to 2021 but I will narrow it down to 1980 to 2021.
How?	I will show three main graphs/charts to convey the data.

Now let's **define our goal**. In as few words as possible, state the goal of your project. In other words, what is the main message your dashboard will be communicating to its audience?

The main point of communication is comparing the winners with the average rating. This visual will help us see the difference between viewer opinions and the academy opinions

Write at least three questions your dashboard will answer. These questions can help you decide which visuals to use later. Each question should also help to achieve the

dashboard's goal. Feel free to take inspiration (or even borrow from) the guiding question listed below the data set's description.

What is the top rated film per year?
And what film won that year?
What is the difference in rating between films that won and the highest rated nominated film per year?

Task 4: Building Your Dashboard

It's finally time to start building your dashboard! Again, you may use either Tableau or Power BI for your project. Before you begin creating the dashboard, review the checklist of requirements and the rubric by which you will be graded. The rubric can be found on the last page of this document.

Below is the list of requirements for your dashboard:

- The dashboard must contain at least three data visualizations. ● The dashboard must contain at least one measure created by the learner. ● The dashboard must be presented in a visually compelling and clear manner.

In the next module, you will have the opportunity to present your dashboard to a small group of classmates before your final presentation. Your dashboard does not need to be complete before module 9.10, but you should have enough to present to your group.

Task 5: Peer Feedback Review

After the peer review is completed, be sure to send the other learners in your group your feedback (and make sure they send theirs).

Take time to review the peer feedback and answer the following questions:

Did your audience understand the goal of your dashboard? If not, why do you think that is the case?

They did not understand because in my first draft my three visuals did not convey one cohesive point or story.

Does your dashboard fulfill all the requirements? If not, what requirements are not fulfilled?

The dashboard was able to convey the rating of nominated vs winning films but was not able to show the major difference between the highest rated and winning film.

What did the audience like about your dashboard?

They liked the temperature chart that helps visualize the rating difference

What did your audience think could be improved about your dashboard?

Add different information to get my main point across.

What changes do you plan on making to your dashboard before the final presentation?

I want to make a barbell graph to better show the viewer the ratings and the difference of opinion between the academy decisions and the audience opinions.

Task 6: Dashboard Improvement

After reviewing your peer feedback, you should have a few things in your dashboard that could be tweaked or improved. Make any final changes to the dashboard so it is ready to be presented during module 9.11.

Task 7: Presenting

During module 9.11, you will present your final dashboard to the class. Again, the dashboard will be graded both by its content and your ability to present it.

We know public speaking can be intimidating. Our best advice is to practice beforehand. Present your dashboard to your pet, roommate, spouse, child, friend, or anyone else willing to listen.

You will have 4 minutes to present the dashboard and an additional 2 minutes for questions.

What should you talk about for those 4 minutes? It's up to you, but the following outline may be helpful:

- Introduce yourself and the data set you selected.
- State the goal of your dashboard and what you are trying to communicate.
- Describe your chosen visualizations and what they represent.
- Share any other details about your dashboard that could be interesting to the audience.

Grading Rubric

Below is the rubric your instructor will use to grade your final project. There are 15 points total. To pass the final project, you must exceed expectations (10/15 points) on the rubric.

	0: No Submission	1: Does Not Meet Expectations	2: Meets Expectations	3: Exceeds Expectations
Data Visualizations	The dashboard's visualizations are incomplete or missing.	The dashboard contains at least one visualization but is otherwise incomplete.	The dashboard contains at least two data visualizations but may be missing other elements.	The dashboard contains at least three well-made visualizations.
Measures	The dashboard contains no learner-created measures.	The dashboard contains a measure that is incorrect or incomplete.	The dashboard contains a measure that does not contribute to the dashboard's goal.	The dashboard contains a measure that contributes to the dashboard's goal.
Presentation	The learner did not present a dashboard.	The learner attempted to present a dashboard but was not able to explain it in a compelling manner.	The learner is sometimes able to explain their dashboard.	The learner presents their dashboard in a clear and compelling way.
Visual Experience	The dashboard is confusing and not visually appealing.	The dashboard contains limited design effort.	The dashboard contains some design effort but may lack clarity.	The dashboard is visually appealing and easy to understand.
Overall Experience	The dashboard does not communicate its defined goal.	The dashboard attempts to communicate its defined goal.	The dashboard somewhat communicates its defined goal.	The dashboard fully communicates its defined goal.