

Response Summary:

Acquire Worksheet

Goal: Identify appropriate data sources, analyze the data, identify data types, variables, list assumptions about the data

Objectives: Students will identify and acquire data from appropriate data sources

Outcomes: Data for the current visualization challenge

1. Student Information *

First Name	Matthew
Last Name	Gallagher
Course (e.g. CGT 270-001)	270-009
Term (e.g. F2019)	F2021

2. Email Address *

gallag80@purdue.edu

3. Visualization Assignment *

- Training Data

Generate

4. Identify appropriate data sources: is the data publicly available? What search methods were used? *

Data source 1	1999-2013 data spread sheet.
Data source 2	College Games Played and Win Percentage.
Data source 3	Combine Data

5. Data format: what format is the data in? Structured vs instructed? All text, a combination, multiple sources? Is it primary or secondary data? *

The data format is structured because the information is displayed very clearly and is quantative. There are fixed columns that include names and numbers. The data is very well organized as well.

6. Data types: what types of data are in the data? How are they stored? What is the access to the data (API, JSON, txt, csv, etc.)? What structure holds the data (data base, spreadsheet, etc.)? *

There is character data because there are names of NFL players. I also

Evaluate

7. Variables: list the data variables? What are the parameters? Give them names. What are the dependent variables and independent variables? *

The data has both independent and dependent variables. Some independent variables include: games player, college played for, college wins, and NFL combine stats. Dependent variables include: Attempts, Yards, Touchdowns, and Fantasy Points. One relationship is that the more games the player had the more yards they would get. Another example is if a football player played for a great college, their wins would be higher than someone that played for a worse college.

8. Audience & Assumptions: list any assumptions you have about the data. Who is your audience? *

I assume that the data presented is fully accurate and credible. My audience would be anyone that has interest in football. Another audience would be analysts and they don't have to be a big fan of sports.

Generate

9. What real life behavior does the data reflect? Does it show patterns of activity, regularity of events, a timeline, population data, etc? Explain. *

The real life behavior is the activity of players when playing football. The data presented contains statistics about what the player did during either a game or a draft combine.

11. What are the weaknesses of the data source? Is it likely that the source will be available in the future? Is the data complete? What is the quality of the data? Is it specific to your needs for the current project? Is the data in the format you need? Are there missing data? Explain. *

This source is reliable for now. Some players on this list are still active in today's league. They could have a breakout year and significantly improve their stats. The data is complete because it covers every major stat that a football player would have. The data format is very easy to understand

12. What information is emphasized? What is the central focus of the data? Explain. *

The main focus are the stats from games because they take up a lot of the chart.

13. At what level of granularity is the data provided? Is the data summarized, or do you have access to the raw data? Is the data categorized or is the data in a format that allows you to create your own categories, etc. Explain. *

I have access to the raw data. All the data is very easy to understand and is displayed clearly. This is a secondary source that took a bunch of information from primary sources.

14. What is the scope of the data? What topics can be covered using the data? Is there a time range/frame? Is the data for a specific area/discipline/demographic etc.? Explain. *

The time range is 1999-2013. The data is targeted towards a variety of NFL players. Some of them are legends and others are names no one has heard of .
