

Response Summary:

1. Student Information *

First Name	Grace
Last Name	Combs
Major	Web Development
Course (e.g. CGT 270-001)	CGT 270-003
Term (e.g. F2019)	SP2022

2. Email Address *

(University Email Address is required.)

gcombs@purdue.edu

3. Visualization Assignment *

- Lab Assignment

Analyze

4. Basic Descriptors: for each data component from the Parse Worksheet, identify basic descriptors (basic statistics). Explain *

pclass: 1, 2, or 3

survived: 1 = yes 0 = no

name: Last name, First name with gender identifier. If woman is married her first name is her husband's and her name is in parentheses. There are 1309 names.

sex: female or male. 466 females, 848 males.

age: number with decimal for infants

sibsp: 1 = yes 0 = no

parch: 1 = yes 0 = no

ticket: 5 or 6 numbers some have letters in front

fare: money with decimal

cabin: letter with room number, some people have more than 1 room

embarked: S, C, or Q for place they embarked

boat: number or letter of lifeboat

body: the number in order of the body being found

home.dest: location of destination

5. Categorize: consider what is similar and what is different? Categorize the data. Are the variables categorical (normal, ordinal, or rank). Are they quantitative (discrete or continuous)? Show categories. Explain. *

pclass: nominal
survived: nominal
name: nominal
sex: nominal
age: ratio
sibsp: nominal
parch: nominal
ticket: nominal
fare: ratio
cabin: nominal
embarked: nominal
boat: nominal
body: ordinal (discrete)
home.dest: nominal

6. Temporal: is the data streaming data? How is it stored (all at one time, over several years in years, days, minutes, seconds)? Explain. *

No the data is not streaming and is not being updated. It is stored all at one time for the single Titanic event.

7. Range and Distribution: what is the distribution of the data? Few values, small size, evenly spread, sparse or dense? Explain. *

The distribution of all the data is quite spread out because the variables represent lots of different things like money, tickets, or only 1s and 0s for some things.

Evaluate

8. Questions and Assumptions: list at least 3 questions you plan to answer with the data or list the questions if they were provided. Must be complete sentences and end in a question mark. What assumptions are you making? *

Question 1	What is the gender ratio of surviving and non-surviving passengers?
Question 2	How many passengers were elderly or very young, what is the survival rate of these groups?
Question 3	Is there a difference in survival rate between 1st, 2nd, and 3rd class passengers?
Assumptions	More men will have survived because there are overall more male passengers. The survival rate of children and the elderly will be low.