

URL to view Results [\[Click Here\]](#)

Response Summary:

Parse Worksheet

Goal: to understand the structure of the data

Objectives: Students will change data into a format that tags each part of the data with its intended use

Outcomes: Every element of the data will be broken into its individual parts

1. Student Information *

First Name	Gabby
Last Name	Willard
Course (e.g. CGT 270-001)	CGT 270-LC4
Term (e.g. F2019)	Fall 2021

2. Email Address *

gwillard@purdue.edu

3. Visualization Assignment *

- Lab Assignment

Understand

4. Parse Data: List each field and its data type. Refer to Fry (page 8-9, 2007) for examples of description of different data types (string, float, character, integer), you can also create user defined types (some combination that uniquely identifies data like the Index type in the Fry 2007 page 9 example) *

For the Titanic Passenger List data set, each field and data type is: pclass (passenger class) is an integer, survived is an integer, name is a string, sex is a string, age is mostly integers but some are floats, sibsp (number of siblings/spouses aboard) is an integer, parch (number of parents/children aboard) is an integer, ticket is mostly integers but some are strings mixed with integers, fare is mostly floats but some are integers, cabin is made up of strings mixed with integers, embarked is a character, boat is mostly integers but some are characters, body is an integer, and home.dest (home destination) is a string.

I would split the name variable into first_name and last_name as their own variables. For the home.dest (home destination), I would split it up into home_country and home_city.

5. Assumptions: List any assumptions you are making about the data and/or the visualization challenge (aka the project) *

For the survived variable, the only values are either 0 or 1 for each record. I am unsure of which value means survived and which value means didn't survive. I am also unsure of what the integer corresponding to the body variable actually means. The embarked variable consists of the characters S and C which I am not sure what those two characters represent.
