

[Download as PDF](#)

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	Zhenghao
Last Name	Fang
Course (e.g. CGT 270-001)	CGT 270-009
Term (e.g. F2019)	F2021

2. Email Address *

fang254@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) *

Most profitable Hollywoods Movie

https://sit.instructure.com/courses/22632/files/2992196?module_item_id=462843

The data is secondary and structured.

There are 471 record and 19 variables

Data Types:

Character: Film, Lead studio, Story, Genre

Float: Domestic Gross (\$m), Foreign Gross (\$m), Worldwide Gross (\$m), Budget (\$m), Profitability (% of Budget), Net Profit (\$m), Opening Weekend (\$m).
Integer: Rotten Tomatoes Audience score, Number of Theatres in US Opening Weekend, Box Office Average per US Cinema (Opening Weekend) (\$)

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

For this dataset, it classify the Story too specific. Some of them could merge into one story type.

The way of counting the profitability is different from each dataset I found. I need to calculate in the same format and unit.

The unit of foreign profit gross is dollar.
