MATH-156 - Exam #1 Sample Questions

Directions:

- You may use a regular or scientific calculator on this exam, but not anything that allows you to store data or create graphs
  - Some calculators are available, but I suggest bringing something you're comfortable with if you have access to one
- Answer each question clearly and succinctly while limiting yourself to the amount of writing specified in parentheses at the end of the question
  - For example, Question #1: Briefly explain . . . (2 sentences)
  - Do not try to get around this limit by using long, run-on sentences
- Do not add unnecessary or superflous information to your answers, this detracts from the clarity of your statements and increases the likelihood that you make an inccorect statement (which you'll be penalized for doing)

Conceptual Questions (20 pts here, expect 40ish pts on the real exam)

Question #1: (10 pts)

For each of the scenarios below, indicate whether summary measure used is *reasonable/useful/justified* or *unreasonable/impossible/improper* (use "Y" to denote the former and "N" to denote the later, no explanations are necessary).

A) The *correlation coefficient* is used to summarize the *relationship* between *birth weight (ounces)* and *length (inches)* in newborn babies
B) A *difference in means* is used to summarize the *relationship* between the *advertising spending (dollars)* and *sales (dollars)* of different start-up companies
C) The *interquartile range* is used to summarize the *spread* of students' *scores (points out 100)* on a math exam
D) The *median* is used to summarize the *geographic region* of all accredited colleges in the United States
E) *Regression* is used to summarize the *relationship* between whether a college is *public* and its *tuition (dollars)*

Question #2: (10 pts)

To evaluate the possible association between stretching and rates of injury, researchers recruit 400 sedintary (non-excercising) volunteers from the local community using an online sign-up link posted on numerous social media platforms. They randomly assign 200 of them to an exercise routine involving stretching and running, and the remaining 200 to an exercise routine involving only running. The researchers found that participants in both groups were equally likely to suffer an injury during the duration of the study. Based upon this information, indicate whether the following statements are true or false (use "T" and "F", no explanations are necessary)

A) There are 400 cases and 2 variables described in this scenario
B) The explanatory variable is whether or not an injury occurred
C) The observed outcome is likely explained by the presence of a confounding variable
D) The researchers chose their participants using a simple random sample
E) This study provides definitive evidence that stretching does not benefit athletes

Application #1 (30 pts, expect there to be two applications of this length on the real exam)

Each year the Academy Awards chooses one movie as the "Best Picture" of that year. This application involves data on the "Best Picture" winning films from 1927 to 2014.

The table below shows genres of these films:

Genre	Action	Adventure	Biography	Comedy	Crime	Drama	Musical	Western
Frequency	2	8	16	14	8	36	2	1

The graph below displays the durations of these films by genre:

Fig 1: Duration of "Best Picture" Winners by Genre (1927–2014)
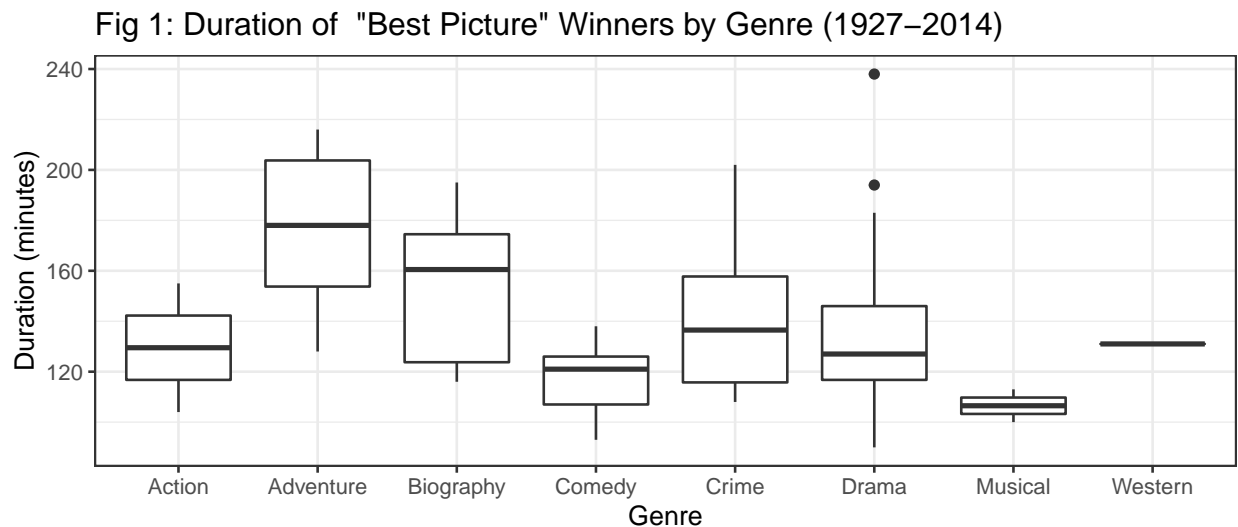
The table below displays the relationship between genre and the month that a film was released:

	Released in the fall/winter/spring	Released in the summer	Total
Action	2	0	2
Adventure	8	0	8
Biography	16	0	16
Comedy	13	1	14
Crime	7	1	8
Drama	29	7	36
Musical	1	1	2
Western	0	1	1
Total	76	11	87

Question #1 (4 pts)

What is a "case" in this dataset? How many cases are there? (one sentence)

Question #2 (4 pts)

Could the correlation coefficient be used to summarize the relationship between the variables shown in Fig 1? Briefly explain (one sentence).

Question #3 (4 pts)

In these data, which genre of films appears to have the longest duration? Briefly explain (one sentence)

Question #4 (4 pts)

In these data, it accurate to say that 75% of biographies were shorter in duration than the median adventure film? Briefly explain (one or two sentences)

Question #5 (7 pts)

Among "Best Picture" winners, does there appear to be an association between genre and whether the film was released in the summer?

Question #6 (7 pts)

In this dataset, 12.6% of films were released during the summers months. Is it reasonable estimate that around 12.6% of movies released in 2021 will premire during the summer months? Briefly explain (three sentences).