# COMP 370 Homework 10 – Data Annotation & Analysis

Assigned Nov 8, 2024  
Due Nov 15, 2024 @ 11:59 PM

## Conceptual Questions

1. When dealing with human coded data, it can be tempting to throw out data items that did not have perfect or majority agreement. Why is this a bad practice?
2. Consider the following analysis tasks. Rank them by the degree to which they are fishing expeditions. Explain your ranking.
   1. Measure the correlation between outside temperature and the amount of foot traffic on St. Laurent on a Saturday night.
   2. Identify the kinds of locations where people go on St. Laurent.
   3. Identify the dietary preferences of people who shop on St. Laurent.
3. For the most extreme fishing expedition identified in #1, explain how more work in the data annotation phase would avoid/limit the open-ness of the question.
4. Diagram and explain how the desire to use automated data annotation can lead to a secondary data science project.
5. For each question below, identify the type of answer required (definition, process, or motivation) and give a precise answer for a young child audience (say somewhere around 5-6 years old).
   1. What is an airplane?
   2. Why do we have stoplights?
   3. When is it appropriate to yell “Fire” in a crowded room?
   4. How do you eat an apple?

## Submission Instructions

Submit a single file called answers.md that contains your answers to the questions above.