

CSC 233 Intro to Data Analytics - Discussion questions from throughout this course (so far)

2026-02-24

Note that this is NOT an exhaustive list of things to study for the exam.

You should ALSO review your HW and project progress to be able to answer questions about, and describe solutions to, hypothetical problems/data analytic tasks

Although I've included weeks when each question was posed/discussed to help point toward which of your notes you may want to review, each topic spans multiple weeks, classroom discussions and assignments.

Reprints of questions posed in this class:

- Remembering Fermi questions - why should we estimate expected results?
 - focus during week 1
- Describe a data analytics workflow to a domain expert, what are the general steps and their purposes?
 - focus during weeks 1-2
- What qualities and practices maximize readability in data analytics work, like your homeworks?
 - focus during HW, peer review
- What are the benefits/tradeoffs of using Lists/Numpy/Pandas classes?
 - focus during weeks 3-4
- How might you handle missing values and how could your choice influence results?

- focus during week 4
- Thinking about combining or grouping data, why do we need to check the results of operations?
 - focus during weeks 4-5
- General steps to data processing/cleaning?
 - focus during weeks 4-5
- Why do we need to “tell stories” - narratives - in data analytics? (provide examples)
 - focus during HW and projects, peer review
- What should you consider in assessing data source quality and what are examples of high quality sources?
 - focus during week 6
- Why do we need to use visualizations in data analytics? (provide examples)
 - focus during weeks 6-7
- What are things to be wary of when drawing conclusions?
 - focus during weeks 6-8