

Closing Remarks

Prof. Sharon McNicholas

STATS 780/CSE 780

1

Introduction

- Over the past few weeks, we have covered a lot of material.
- No topic was covered exhaustively; in fact, much of the coverage was far from exhaustive.
- The intention was to give a broad coverage of some of the relevant material, and I think we did that.
- I want to finish by returning to a few recurring themes.

2

Recurring Themes

- The data matter, at all stages.
- Analytic decisions often need to be made in the context of the data — during the past few weeks, I have often said things such as: “here is where we would talk to the data person/collaborator”.
- The goal/purpose is very important, e.g., modelling versus prediction.
- Familiarity (and competence) in software, e.g., R, is very important.
- Data treatment is key, e.g.:
 - Are variables coded correctly?
 - Do we need to scale the data?
 - Etc.

3

Other Topics

- Despite the fact that we covered a lot of topics, there is much more out there (in the data universe!).
- This course focused on some topics in data analytics but data warehousing/storage is also extremely important.
- Data cleaning is another important and interesting topic; in fact, some interesting work is ongoing at Mac.

4

Making Yourself A Better Data Scientist

- If you like, for Data Scientist, you can read “Statistician” or “Computer Scientist” or “Software Engineer”, etc.
- Maybe you are thinking along the lines: what can I do to make myself a better data scientist?
- The answer might be that you need to become a better programmer, or improve your knowledge of statistical analysis, etc. The answer will depend on your current strengths and weaknesses.
- In a rapidly evolving field, I think continual improvement is very important.

5

Final Comments

- I hope the course has been enjoyable and that you found the material interesting.
- I also hope that the mode of delivery was useful.
- Office hours will take place the next two Wednesdays from 1–2pm.
- There will be an interesting guest speaker in class next week, and don't forget that your final projects are due (please read the Project outline and remind yourselves of the due dates/times).

6