

Problem set 2

Alexandra Penner

January 2018

1 Tools of Data Scientist

- Text editor such as LaTeX or Notepad++
- Versioning software such as Git
- Software for data cleaning and Analysis Such as R or Python
- software to manage large amounts of data such as SQL, or spark
- software to access remote computing such as SFTP or SSH