University of Virginia
DS 5559: Big Data Analytics
Final Project Grading Rubric

Final Project (30 points)
Refer to the *Final Project Instructions* for details.
Code, Paper, and Presentation are each worth 10 points.

Within each component, grading will be implemented as follows:

Code (10 points)

The code should be in Pyspark. In specific circumstances, python code may be admissible, but you must discuss this with the instructor first. If small parts of the code are in Python (e.g., reading into a pandas dataframe), this is fine. No other coding languages are admissible.

The code needs to be clearly written and documented in a Jupyter notebook (ipynb format). If you have all of the code in a single cell, this is fine.

For full credit, the code needs to include these sections in a clean, commented, and comprehensive manner. For each component below, unclear or partial work will be given half credit, and no work will be given no credit.

Component	Points
Data import and preprocessing	2
Data splitting/sampling	1
Exploratory data analysis with at least 2 graphs	2
Model construction with at least 3 models	3
Model evaluation	2

Presentation (10 points)

The presentation grading breakdown is shown in the table below.

Component	Points
Content	4
Organization and aesthetics	3
Delivery	3

To earn all points for **content**, the presentation must include all items from the list below. Points will be deducted for missing items.

Component
Executive summary: discuss the research question and what you have found
Data summary: explain the target (response) variable, the predictors, sampling, etc.
Variable transformations and preprocessing
Models constructed (or planned to be constructed)
Model performance
Conclusions and future research

To earn all points for **organization and aesthetics**, the presentation must be clear, easy to follow, and appealing. Slides should be brief (make the audience listen, not read), and graphs should be well-labeled and easy to interpret. The presentation does not need to be fancy.

To earn all points for **delivery**, the speakers need to have command of the material and present with volume and confidence.

Paper (10 points)

For full credit, the project writeup should include the sections below. It should be clear, and only as lengthy as it needs to be (but no more than 10 pages). If a section is unclear, partial credit will be given. If a section is missing, no credit will be given. Do not include a dump of code, tables, values, or figures; only show what is essential.

Section of Paper
Abstract
Data and Methods
Results
Conclusions