MPA 634  
Data Science for Managers  
Midterm I: Winter 2021

Exam parameters:

* Closed book
* No time limit
* Complete in one sitting
* Submit before starting the second part of the exam

# I. Definitions and Concepts

1. Compare and contrast filter, mutate, and summarize. Your answer should include an explanation of how each one affects data frames or tibbles.
2. Identify and explain at least three different ways that you can use a violin plot to determine whether or not a distribution is positively or negatively skewed.
3. Carefully explain how the whiskers of a boxplot are constructed. How do whiskers help us identify outliers?
4. Define and explain each of the seven parts of the grammar of graphics by writing a script that illustrates the definition of each part.
5. **Compare** the location, scale, symmetry, and outliers of city and highway mileage using the following information:

A close up of a map

Description automatically generated

type\_of\_driving Mean Median Standard\_Deviation Interquartile\_Range

*<chr>* *<dbl>* *<dbl>* *<dbl>* *<dbl>*

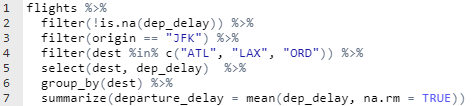
1 cty 16.9 17 4.26 5

2 hwy 23.4 24 5.95 9

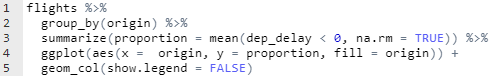
* 1. Location
  2. Scale
  3. Symmetry
  4. Outliers

# II. Line by Line Code Interpretation (Don’t interpret the first line)

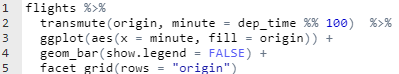
Code Chunk I



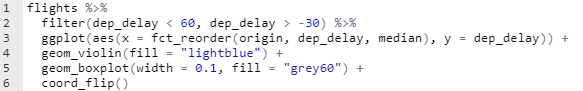
Code Chunk II



Code Chunk III



Code Chunk IV



Code Chunk V

