MPA 634  
Data Science for Managers  
Midterm I: Fall 2019

# I. Definitions and Concepts

1. Compare and contrast hypothesis generation and hypothesis confirmation.

**Hypothesis generation**

**Hypothesis confirmation**

**Key difference**

1. Define and illustrate all seven parts of the grammar of graphics by outlining a script that creates bar graphs from the diamonds data frame. Remember that possible variables in the diamonds data frame are price, carat, cut, color, and clarity.

|  |  |
| --- | --- |
| **Data** |  |
| **Aesthetics**: |  |
| **Geometric Objects**: |  |
| **Stats**: |  |
| **Position**: |  |
| **Coordinate System** |  |
| **Facet**: |  |

1. Use Venn diagrams to define and illustrate the meaning of "&", "|", and "!".
2. Locate Tukey's five number summary on the following boxplot.



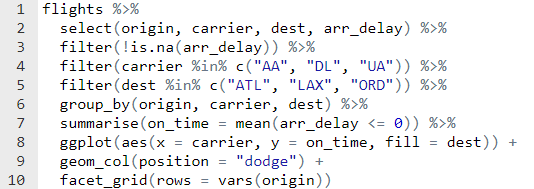
1. Use the following violin plots to compare the shapes of the distributions of highway miles per gallon for suv and subcompact vehicles .



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

# II. Line by Line Code Interpretation

## Code Chunk 1



## Code Chunk II

