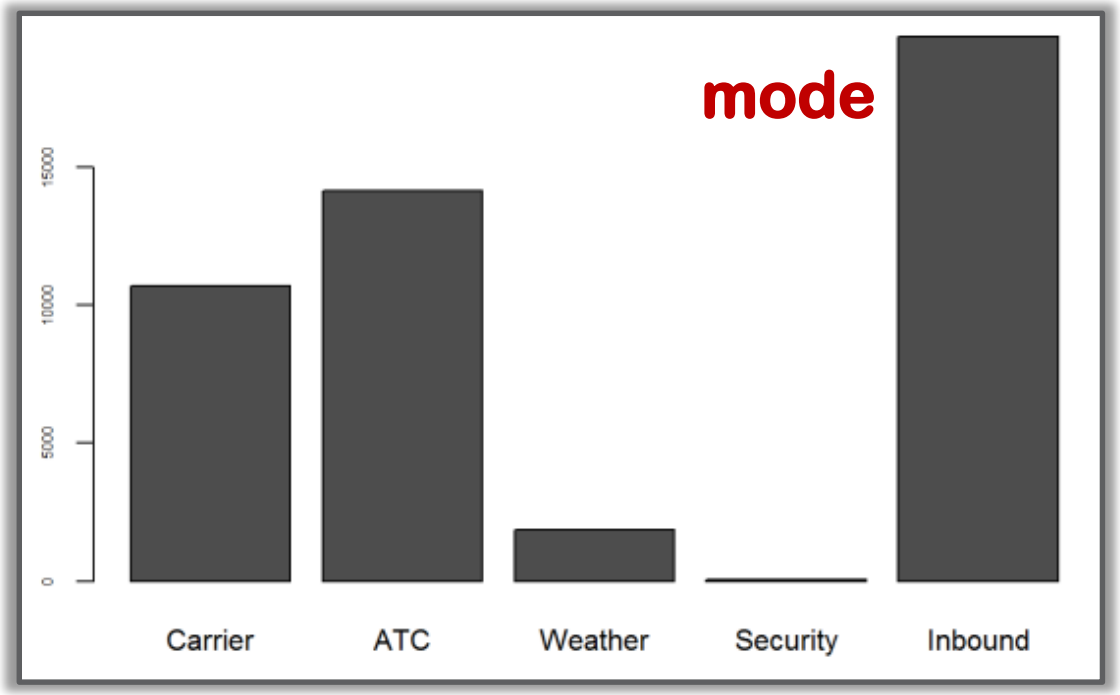
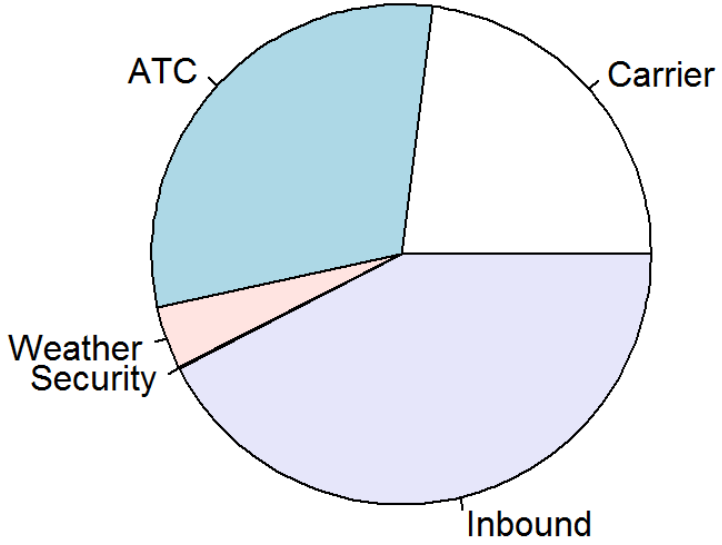
Univariate: Categorical: Mode & Expected Value



- Central tendency: mode & expected value
- Visualization: bar charts & pie charts



Prof. Nagiza F. Samatova
samatova@csc.ncsu.edu
Department of Computer Science
North Carolina State University

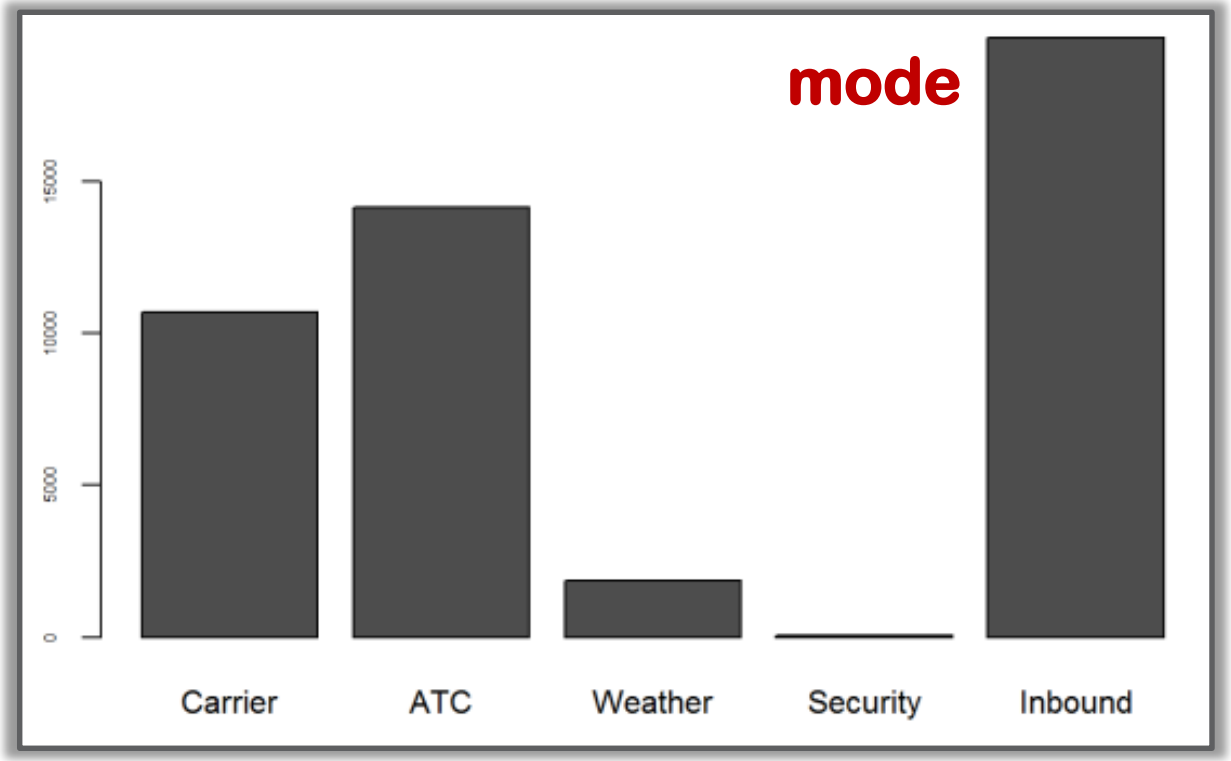
Centrality Tendency: Mode and Expected Value

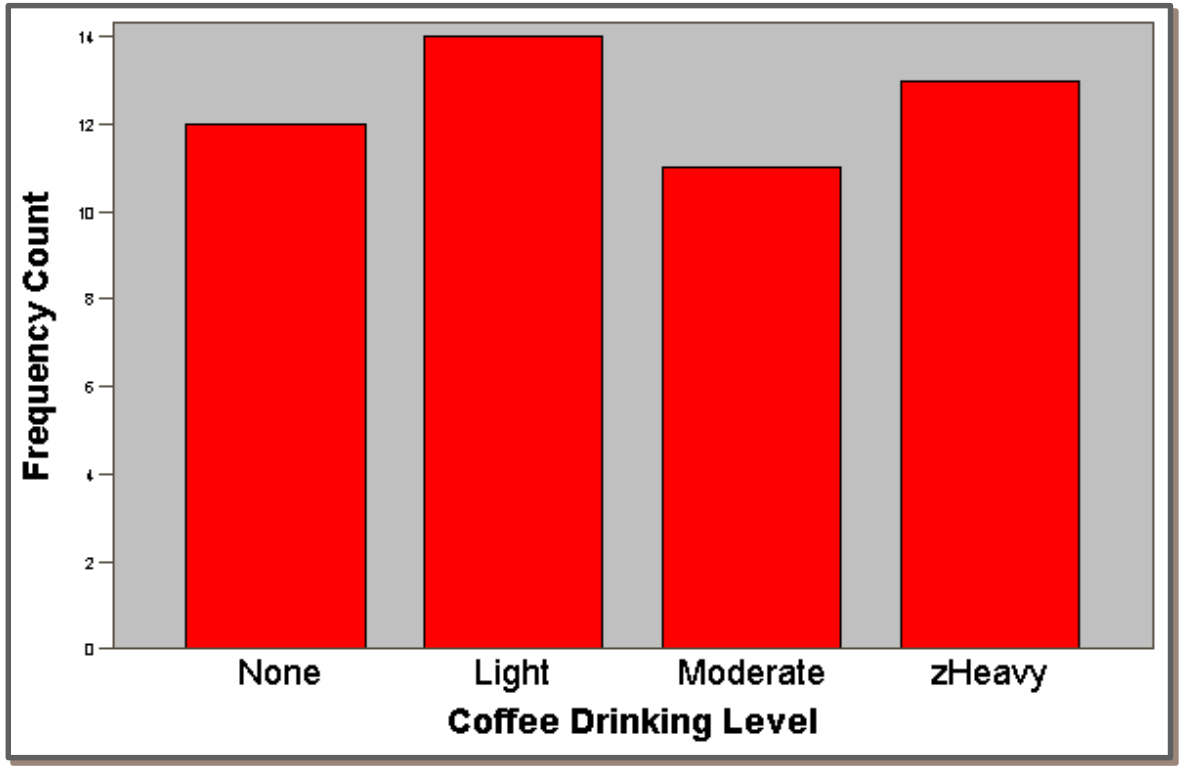
Measure	Description	Comments
Mode	the most frequently observed value in the data	
Expected Value	average value based on categories' probability of occurrence	categories are numeric
Bar charts	frequency of proportion for each category plotted as bars	
Pie charts	frequency of proportion for each category plotted as wedges in a pie	

Bar Chart: Categorical Variables with plt.bar()

Show frequency or proportion in each category

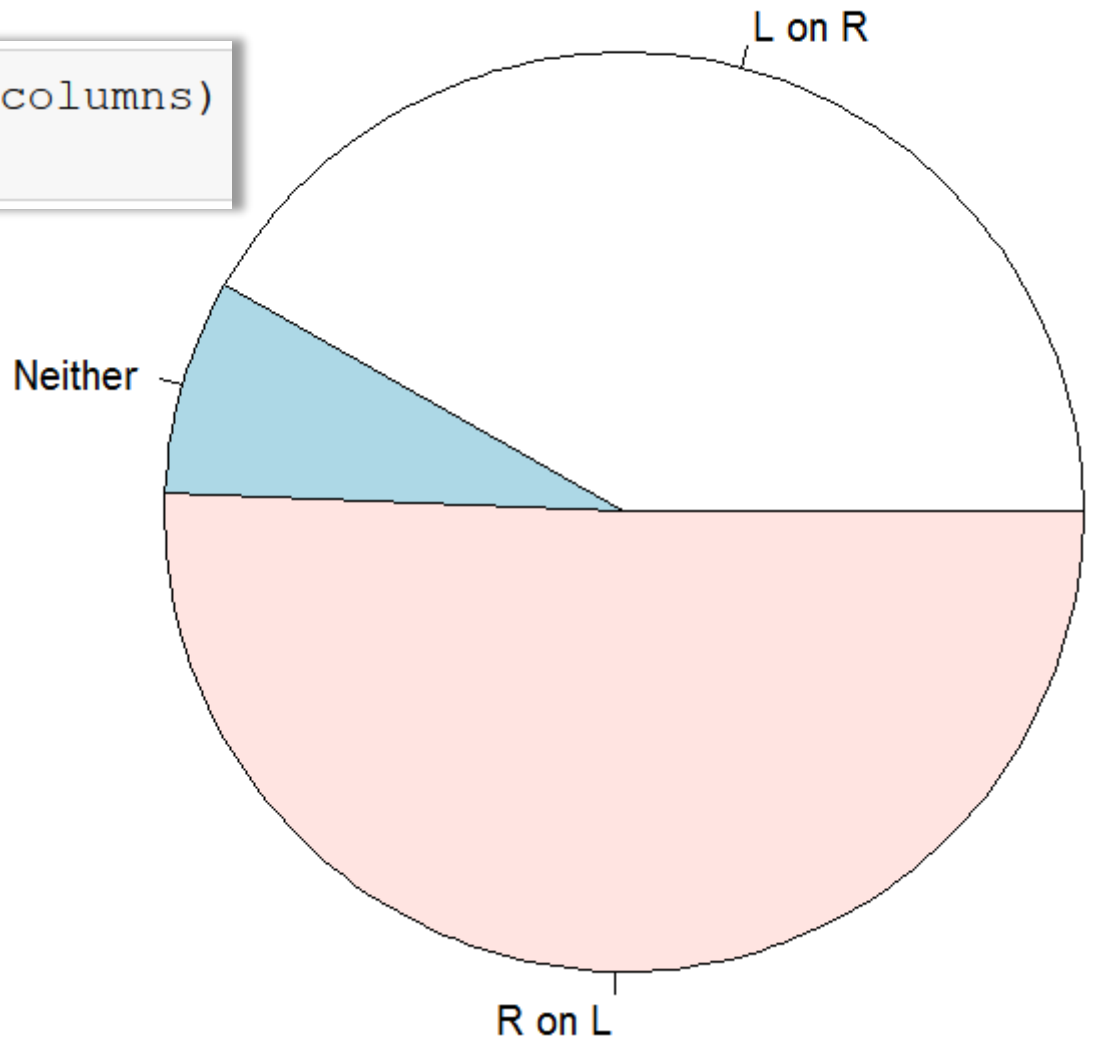
```
1  plt.bar([1,2,3,4,5],dfw.iloc[0])
2  plt.xticks([1,2,3,4,5], dfw.columns)
3  plt.show()
```





Pie Charts

```
1  plt.pie(dfw.iloc[0],labels = dfw.columns)
2  plt.show()
```



Expected Value for Numeric Categorical Variables

- **Discrete (numeric) categorical variables**
  - Expected value of five years of profits from a new acquisition
  - Expected cost saving from new patient management software at a hospital
- **Calculation: Expected Value**
  - Multiply each outcome by its probability of occurring
  - Sum these values
- **Example:**
  - Marketer offers free webinars to generate leads
  - 5% attendees will sign up for $300 service
  - 15% attendees for the $50 service
  - 80% will not sign up for anything
  - Expected value: weighted mean in which weights are probabilities
    - EV = (0.05) (300) + (0.15)(50) + (0.8) (0) = 22.5
    - Expected Value of a webinar attendee is thus $22.50 per month