

# Quantitative Data: Histograms

*Reed Coots*

# What are Quantitative Variables?

Variables that have a numerical value (quantity) that we can perform mathematical operations on

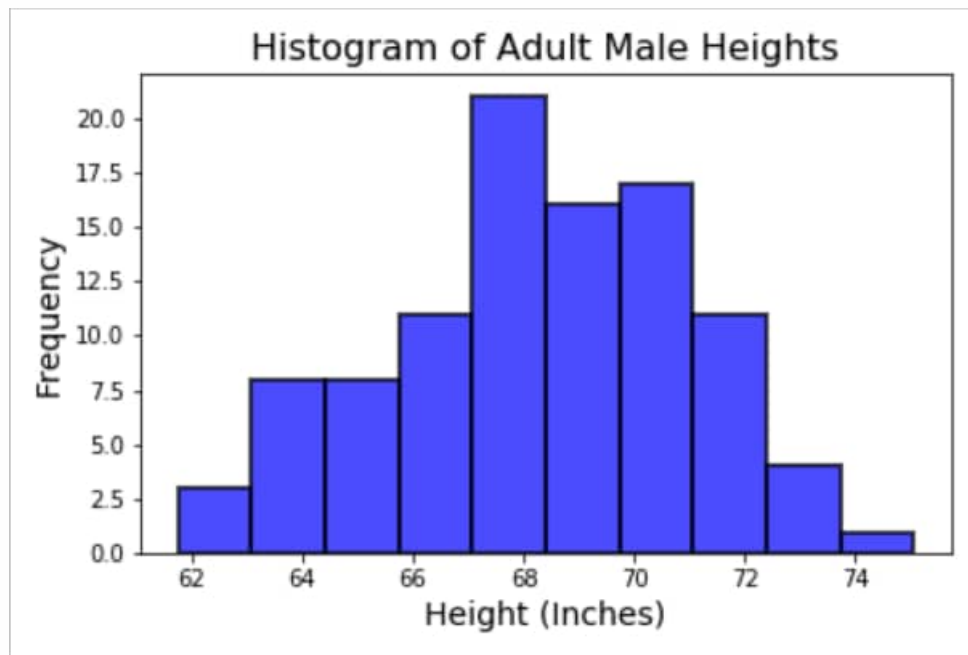
**Examples:** Height, weight, income, test scores, shoe size, number of “heads” after 10 coin flips



# Why Use Histograms?

## Adult Male Heights

66.3
75.1
67.9
67.6
70.0
69.9
64.8
...



## 4 Main Aspects

- Shape -** Overall appearance of histogram. Can be symmetric, bell-shaped, left skewed, right skewed, etc
- Center -** Mean or Median
- Spread -** How far our data spreads. Range, Interquartile Range (IQR), standard deviation, variance.
- Outliers -** Data points that fall far from the bulk of the data

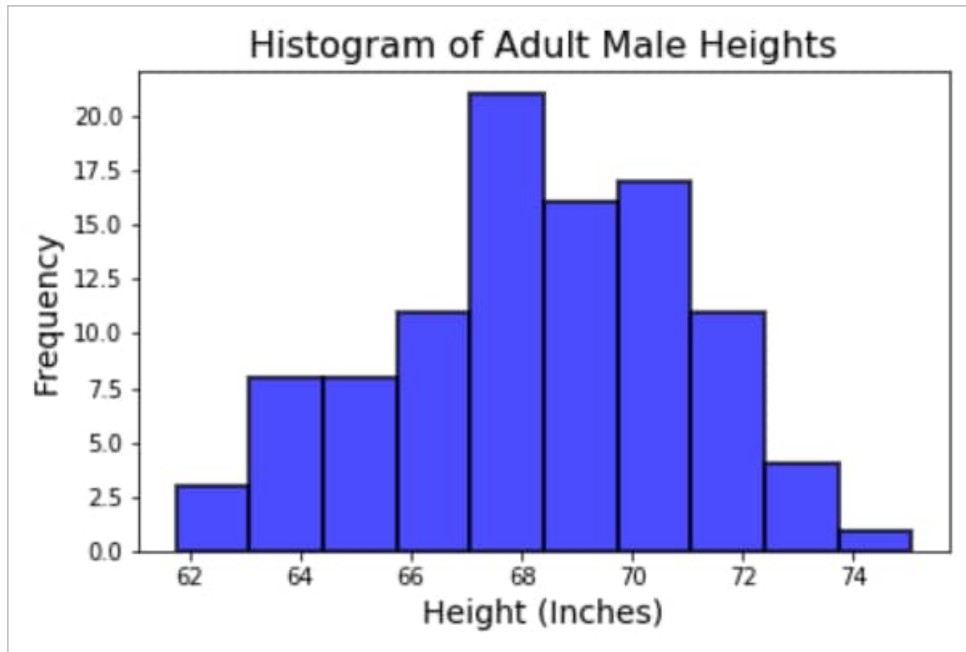
# Adult Male Heights

**Shape**

**Center**

**Spread**

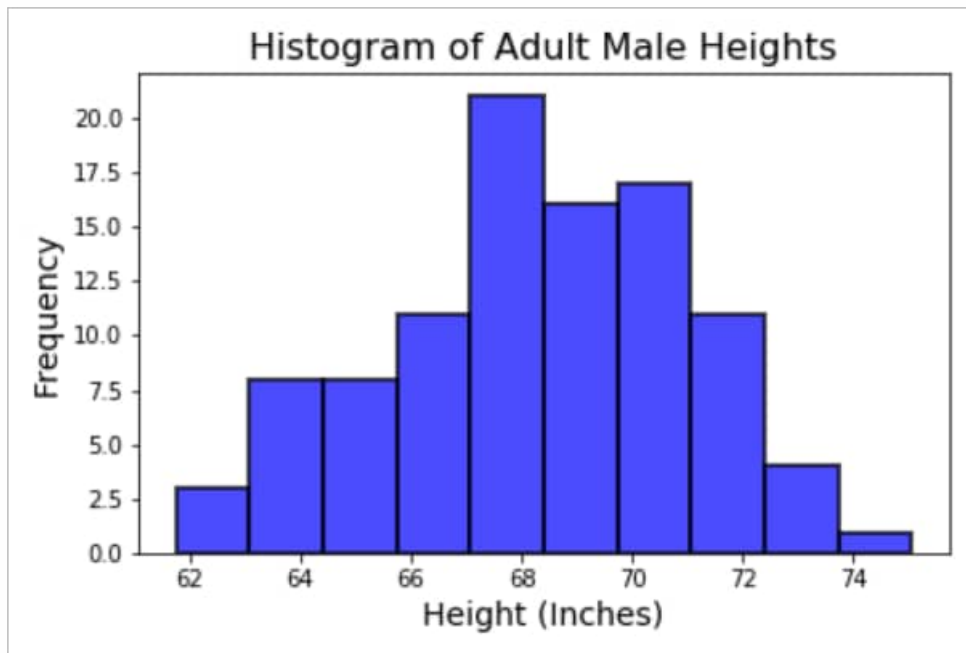
**Outliers**



# Adult Male Heights

## Putting it all together:

The distribution of adult male heights is roughly bell shaped with a center of about 68 inches, a range of 13 inches (62 to 75), and no apparent outliers.



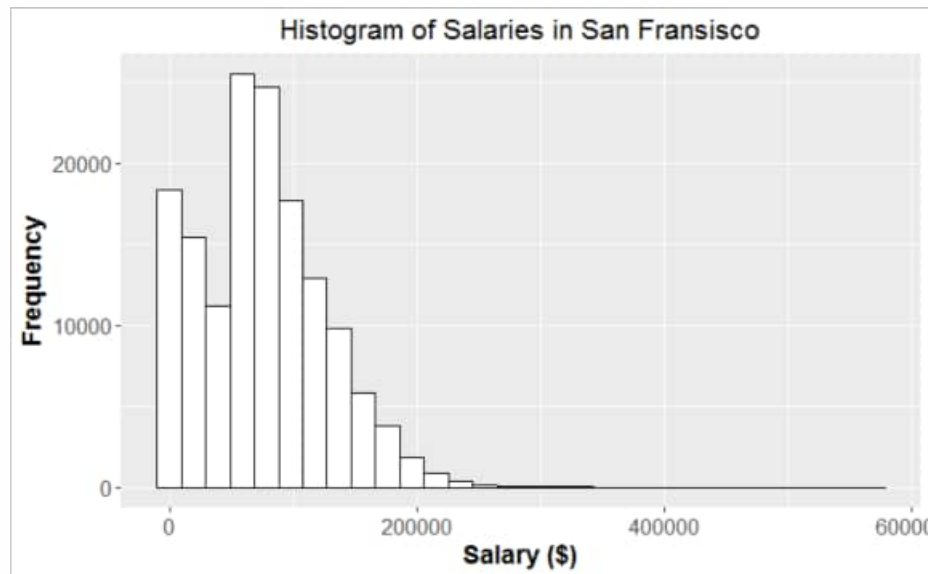
# Salaries in San Francisco (2011-2014)

**Shape**

**Center**

**Spread**

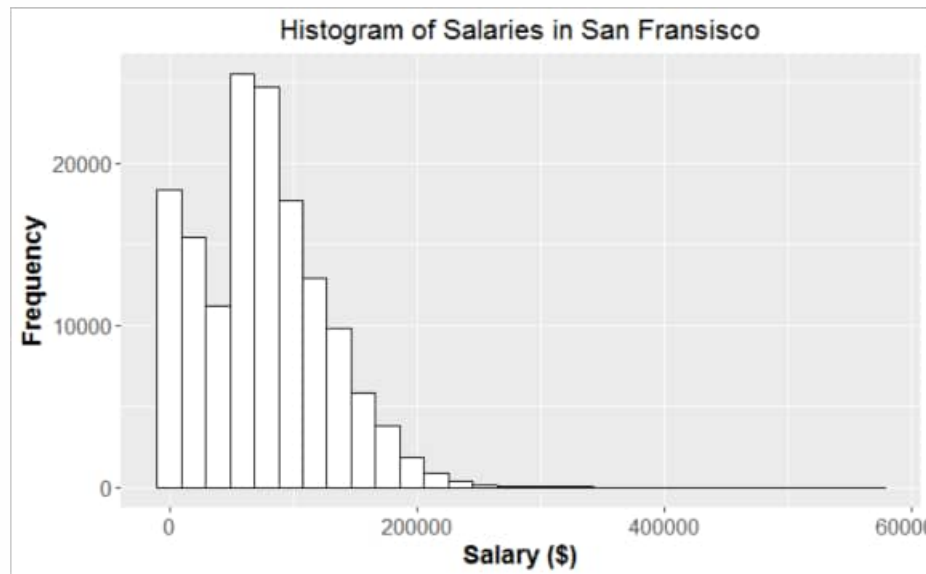
**Outliers**



Source: <https://www.kaggle.com/kaggle/sf-salaries/data>

# Salaries in San Francisco (2011-2014)

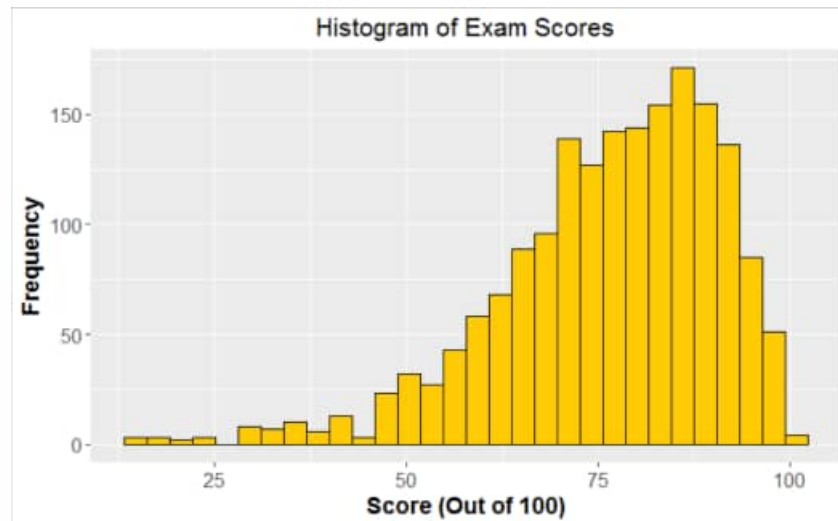
The distribution of salaries in San Francisco is bimodal and skewed to the right, centered at about \$80,000 with most of the data between \$40,000 and \$120,000, a range of roughly \$600,000, and outliers are present on the higher end.



Source: <https://www.kaggle.com/kaggle/sf-salaries/data>

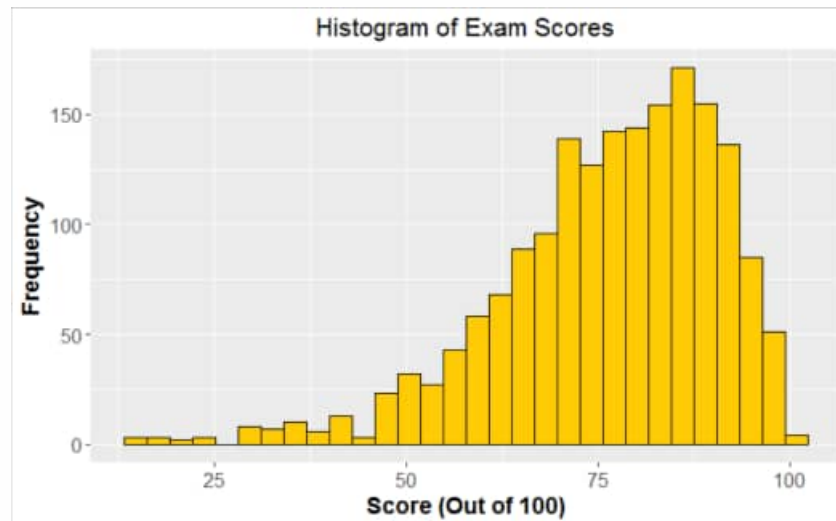


# Exam Scores



# Exam Scores

The distribution of exam scores is skewed left, centered at about 80 points with most scores being between 65 and 90 points, a range of roughly 85, and some outliers are present below 50 points.



# Summary

- Histograms allow us to display data graphically
- 4 main aspects we use to describe the data
  - **Shape**
  - **Center**
  - **Spread**
  - **Outliers**
- Your one sentence summary should allow for any person to read it and have a general understand of what your data looks like

# Attributions

“Tanita-HD-372-bathroom-scale-I” by Thinking Tree is licensed under CC BY-2.0

“Batovka Shoe” by Bata Shoe Museum is licensed under CC BY-SA 3.0