

**Data Appendix**  
**Dataset: Gender Wage Gap Data**

The data file contains five variables, with each row containing information about a different year. The unit of observation is year.

**Variables:**

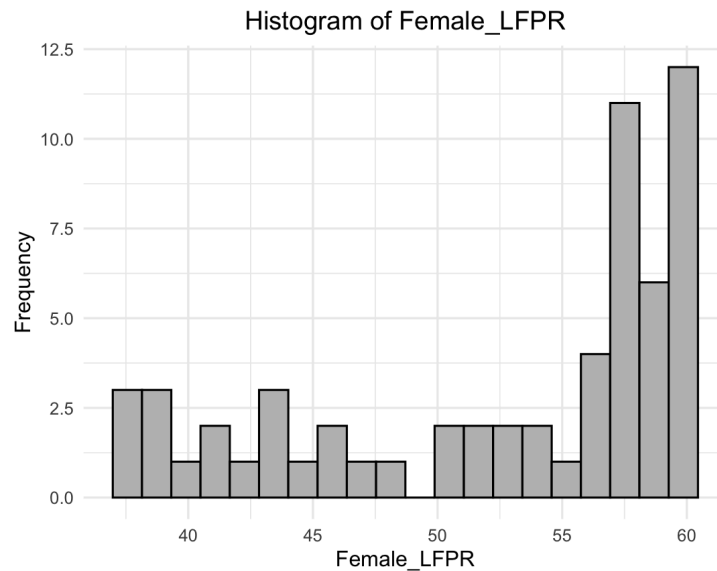
**Year**

- Definition: Year in which data was recorded
- Units of Measurement: N/A
- Exact Name of Variable in original data file: Year
- 60(0)
- Steps for processing the variable for cleaned dataset: No changes were made

**Female\_LFPR**

- Definition: The percentage of US women 16 years or older who participate in the labor force
- Units of Measurement: Percentage
- Exact Name of Variable in original data file: Female\_LFPR
- 60(0)
- Steps for processing the variable for cleaned dataset: No changes were made

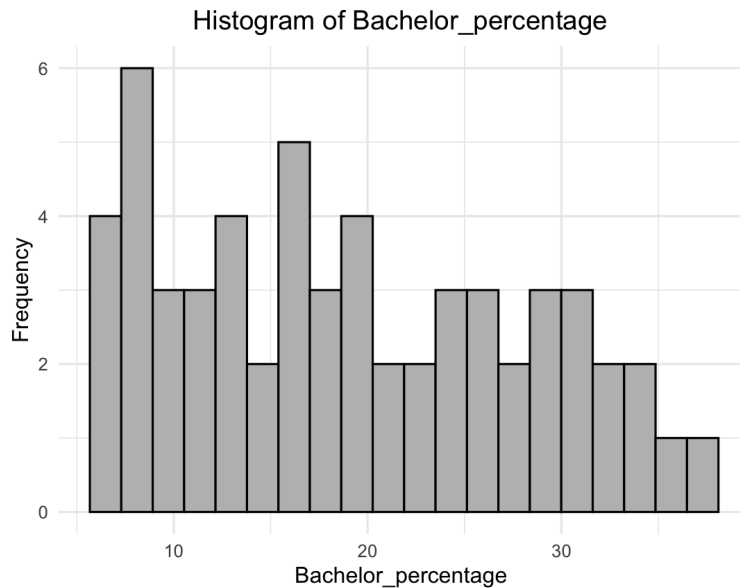
Minimum	25th Percentile	Median	Mean	75th Percentile	Maximum	Standard Deviation
37.70	46.15	56.75	52.63	58.83	60.00	7.509787



### **Bachelor\_percentage**

- Definition: The percentage of US women 25 years or older who have attained at least a bachelor's degree
- Units of Measurement: Percentage
- Exact Name of Variable in original data file: Bachelor\_percentage
- 60(2)
- Steps for processing the variable for cleaned dataset:
  1. Removed rows with N/A values, 1961 and 1963

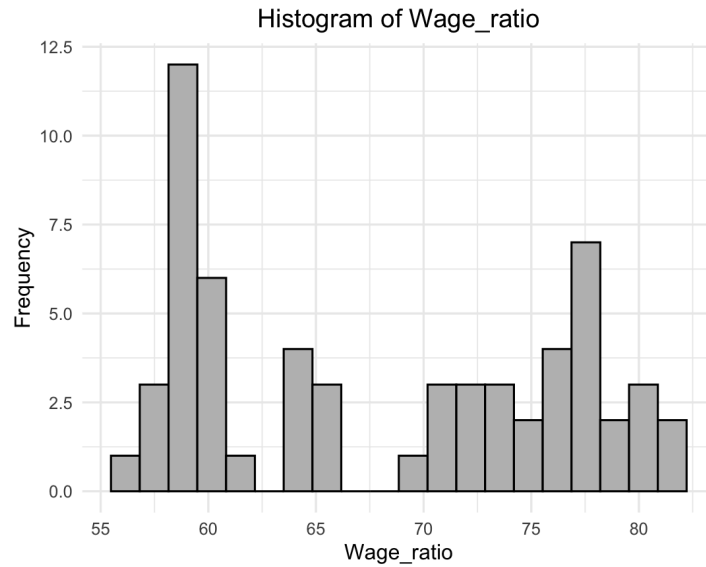
Minimum	25th Percentile	Median	Mean	75th Percentile	Maximum	Standard Deviation
5.80	11.47	18.50	19.16	26.40	36.60	8.938753



### **Wage\_ratio**

- Definition: The ratio, out of 100, of female to male earnings for full-time, year-round workers
- Units of Measurement: N/A
- Exact Name of Variable in original data file: Wage\_ratio
- 60(0)
- Steps for processing the variable for cleaned dataset: No changes were made

Minimum	25th Percentile	Median	Mean	75th Percentile	Maximum	Standard Deviation
56.60	59.40	67.95	68.17	76.60	82.00	8.340806



### **First\_Birth\_Median\_Age**

- Definition: Median age of US women to give birth to their first child
- Units of Measurement: Years
- Exact Name of Variable in original data file: First\_Birth\_Median Age
- 60(0)
- Steps for processing the variable for cleaned dataset: No changes were made.

Minimum	25th Percentile	Median	Mean	75th Percentile	Maximum	Standard Deviation
22.24	22.74	24.31	24.35	25.67	27.41	1.588414

