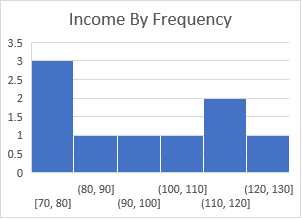
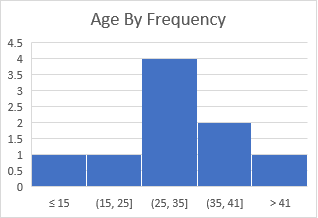
Homework 2



I wanted to choose bin widths that would keep the data clear to visualize. The income data seemed to have no outliers as no use of overflow or underflow bins was necessary. The age data required underflow and overflow bins for its data suggesting it contains outliers.

A graph with blue squares

Description automatically generatedA graph with blue squares

Description automatically generated

I used pandas built in histogram plots for this

1. Age Statistics:

Mean 136.7777778

Standard Deviation 323.9235163

Z-scores:

-0.419166164

-0.3450746

-0.338900303

-0.329638857

-0.32346456

-0.314203115

-0.298767372

-0.295680224

2.664895195

1. mean = df["Age"].mean()

print(mean)

std = df["Age"].std()

print(std)

z\_score = []

for i in df["Age"]:

x = (i - mean) / std

z\_score.append(x)

print(z\_score)

I once again leaned on pandas for this

1. A screenshot of a computer

   Description automatically generated