Jon Organ

CS677

Assignment 6, Store

**Question 1:**

Product data file read...

**Question 2:**

There are 7 categories in the file

**Question 3:**

Clothing has the most subcategories (17), and New Arrivals has the least subcategories (6)

**Question 4:**

Category: Clothing, median price: $89.0, average price: $215.04

Category: Shoes, median price: $130.0, average price: $185.11

Category: Handbags, median price: $698.0, average price: $1017.72

Category: Jewelry, median price: $325.0, average price: $1003.33

Category: Accessories, median price: $198.0, average price: $661.1

Category: Maternity, median price: $68.0, average price: $82.38

Category: New Arrivals, median price: $89.0, average price: $258.01

**Question 5:**

Saving First histogram...

A graph of a number of red rectangular bars

Description automatically generated

Saving Second histogram...

A graph with red bars

Description automatically generated

Saving Third histogram...

A graph with red rectangular bars

Description automatically generated

**Question 6:**

Saving Model\_1 first error histogram...

A graph with blue rectangular bars

Description automatically generated

Saving Model\_1 second error histogram...

A graph with blue bars

Description automatically generated

Saving Model\_1 third error histogram...

A graph with blue bars

Description automatically generated

Saving Model\_2 first error histogram...

A graph with blue bars

Description automatically generated

Saving Model\_2 second error histogram...

A graph with blue rectangular bars

Description automatically generated

Saving Model\_2 third error histogram...

A graph of a number of digits

Description automatically generated

**Question 7:**

For the first digit, the 2 Model was closer to real distribution

Model 1 RMSE: 0.08124630523259022

Model 2 RMSE: 0.024061971648113367

For the first digit, the 1 Model was closer to real distribution

Model 1 RMSE: 0.056836391594678

Model 2 RMSE: 0.11003300700352049

For the first digit, the 1 Model was closer to real distribution

Model 1 RMSE: 0.10136039506146796

Model 2 RMSE: 0.1505157856971232

**Question 8:**

For the first digit:

The Clothing category was the closest to equal weight P

For the second digit:

The Handbags category was the closest to equal weight P

For the third digit:

The Clothing category was the closest to equal weight P

**Question 9:**

For both models it seems they were most accurate with the third digit. The second digit for

both models also seem to be decently inaccurate concerning the 4 through 7. The first model

was very far off in terms of the distribution of ones in the first digit place.