# Programing Languages for Statistics

STAT2005

Concluding Remarks

# R and SAS

We can see both R and SAS have their relative strengths when it comes to statistical programming.

R is open-source, highly flexible and can be easily extended with user-created packages, making it a popular choice for data analysis and visualization.

SAS is known for its ability to handle raw data efficiently, and it has a wide range of built-in functions and procedures for data cleaning, transformation, and analysis.

# A workflow with SAS and R

One could therefore take advantage of both programming languages to tackle data analysis problems in practice.

The simple example below demonstrates this workflow.

1. Use SAS DATA step to read in raw data.
2. Output the SAS dataset into a CSV file.
3. Use R to read in the CSV file as a data frame.
4. Use ggplot2 to draw a graph using that data frame.

# Raw data

```
Item_Name;Item_Price;Number_In_Stock;Number_Sold
Apple;$1.50;50;25
Banana;$1.00;60;40
Grape;$2.00;30;15
Orange;$0.75;80;60
Pineapple;$3.00;20;10
Blueberry;$4.00;15;5
Strawberry;$3.50;40;20
Mango;$2.50;25;15
Lemon;$1.25;100;50
Peach;$2.25;35;20
```

# Use SAS DATA step to read in raw data

```
DATA sales;
     INFILE "C:/sales_data.txt"
     DLM=';' FIRSTOBS=2;
     INPUT
          Item_Name : $20.
          Item_Price : COMMA5.
          Number_In_Stock
          Number_Sold;
RUN;
```

# Output the SAS dataset into a CSV file

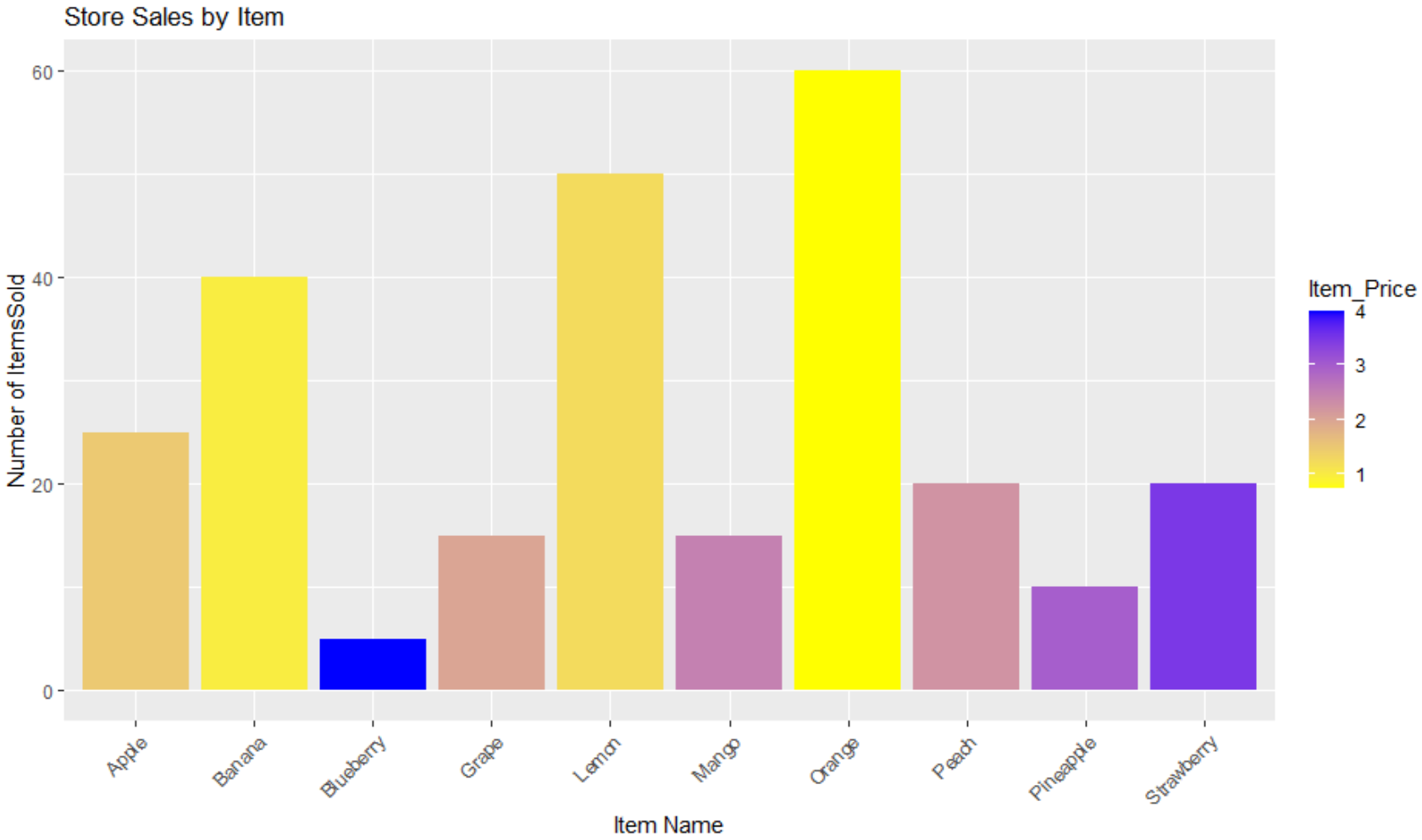
```
PROC EXPORT
     DATA=sales
     OUTFILE="C:/sales_data.csv"
     DBMS=CSV
     REPLACE;
RUN;
```

# Use R to read in the CSV file

```
sales_data <- read.csv('sales_data.csv')
```

# Use ggplot2 to draw a bar plot

```
library(ggplot2)
# Plot the number of items sold by item name
ggplot(sales_data, aes(x=Item_Name,
     y=Number_Sold, fill=Item_Price)) +
geom_bar(stat='identity') +
scale_fill_gradient(low='yellow', high='blue') +
theme(axis.text.x =
          element_text(angle = 45, hjust = 1)) +
labs(x='Item Name', y='Number of ItemsSold',
     title='Store Sales by Item')
```


Store Sales by Item
Number of ItemsSold
60
40
20
0
Apple
Banana
Blueberry
Grape
Lemon
Mango
Orange
Peach
Pineapple
Strawberry
Item Name
Item_Price
4
3
2
1