AD 571: Business Analytics Foundations

LCR Session (Entire Class) Week 4 (5/30/2023)

AGENDA

1. Summary Weeks 1 to 3: Q & A
2. Preparation for Week 4:
   - A. Introduction to Lecture 07 and Lecture 08
   - B. Individual Exercise: Tutorial: Preparation For Assignment 4 - Predictive
     Tutorial: Time Series Analysis in R
     Tutorial: Regression Analysis in R
   - C. Group Discussion Forum W4 & Quiz 4, Exercise M4: Q & A
3. Assignment 4: Preview & Examples (Due: 06/05/2023 at 11:59 pm ET)

2

Preparation for Week 4:

Module 4

May 30 – June 5

Topics:	Lecture 7: Predictive Business Analytics: Basic Concepts and Applications Lecture 8: Model Building for Selected Predictive Business Analytics
Readings:	Lecture 7: • Lecture Notes • Evans, Chapter 8 • Wickham & Grolemund, Chapter 23.3 Lecture 8: • Lecture Notes • Evans, Chapter 9 • Wickham & Grolemund, Chapter 24.3
Tutorials:	Assignment 4 Tutorial: Predictive Analytics Regression Analysis in R Time Series Forecasting
Discussions:	Discussion 4 • Initial post due Thursday, June 1 at 11:59 PM ET. • Respond to at least two of your classmates' posts by Monday, June 5 at 11:59 PM ET.
Assignments:	Assignment 4: Predictive Analytics, due Monday, June 5 at 11:59 PM ET Evans Textbook: Module 4 Exercises, due Monday, June 5 at 11:59 PM ET
Assessments:	Quiz 4, available from Saturday, June 3 at 9:00 AM ET to Monday, June 5 at 11:59 PM ET
Live Classroom:	Module 4 Live Classroom Session — Tuesday, May 30, at 8:00 PM ET Technical Session — Thursday, June 1, at 8:00 PM ET

Preparation for Week 4

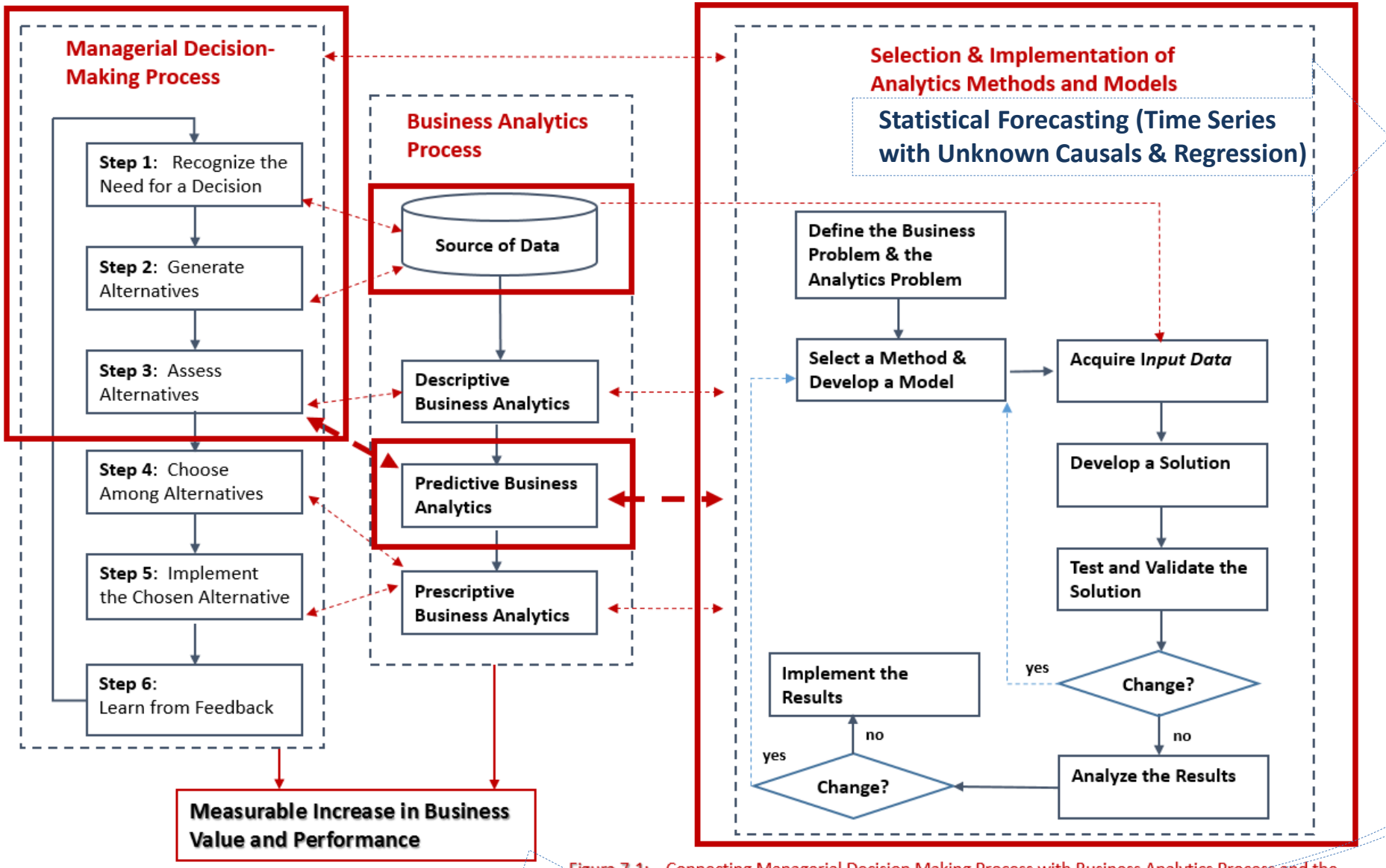


Figure 7.1: Connecting Managerial Decision Making Process with Business Analytics Process and the Implementation of Selected Business Analytics Methods & Models: **Predictive Analytics**

Moving Averages

Single Exponential Smoothing

Double Exponential Smoothing

Triple Exponential Smoothing

Multiple Regression

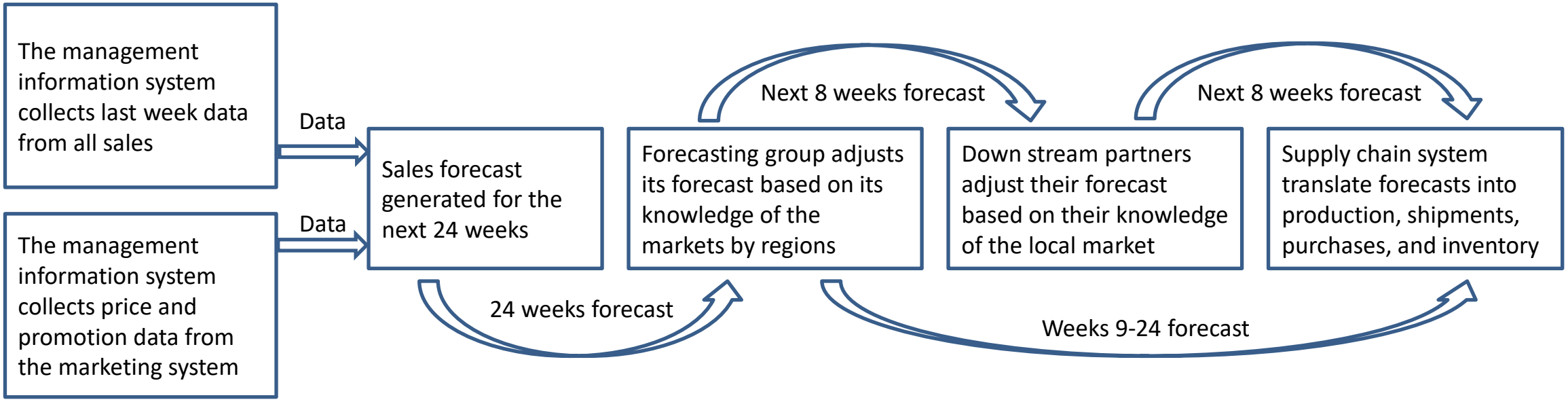
Evaluation of the Results

Recommendations:

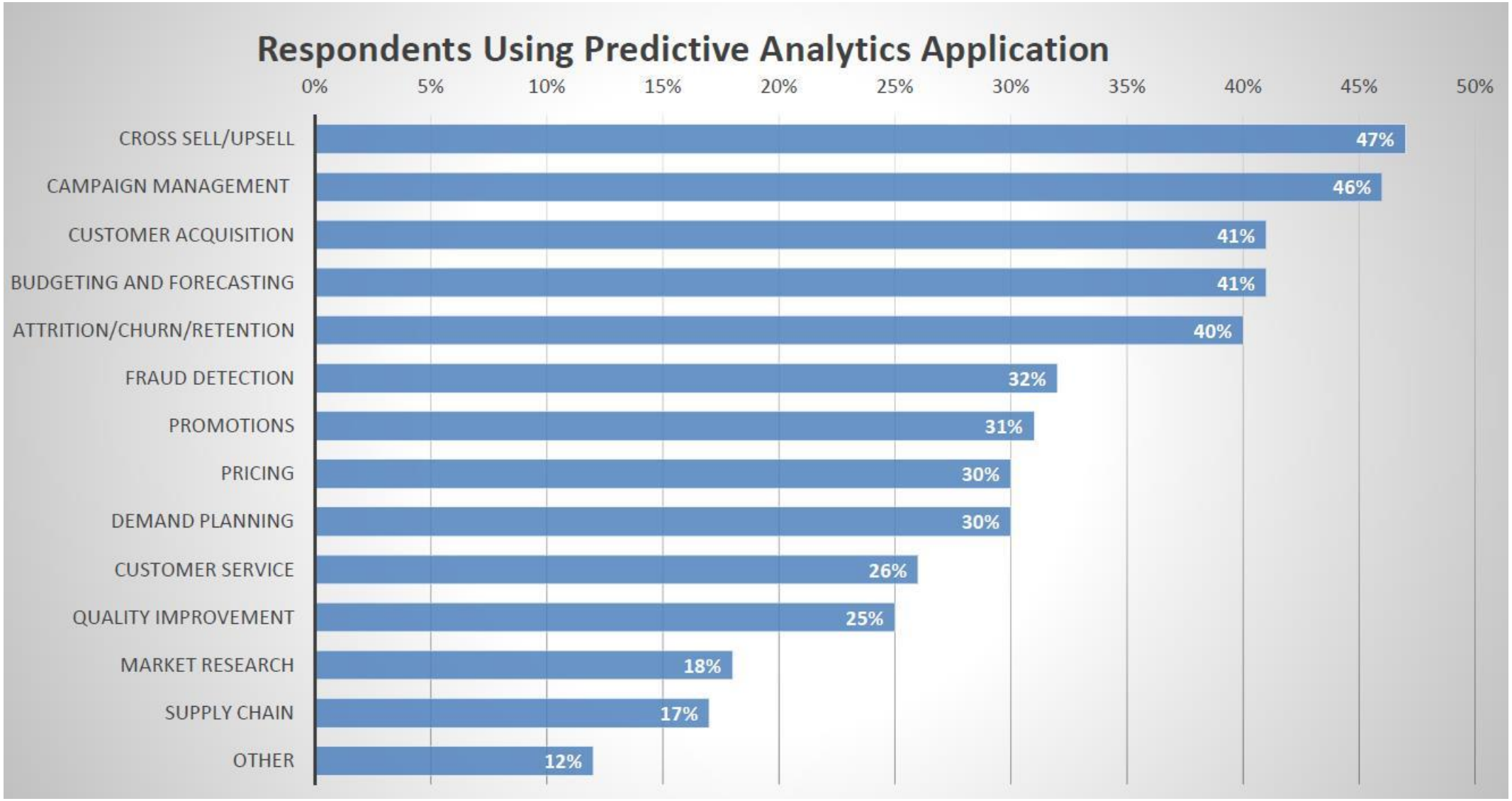
→ Business Analytics Process (overall & next steps)

→ Managerial Decision-Making Process (overall & next steps)

2 A Preparation for Week 4



Preparation for Week 4: Predictive Analytics Applications

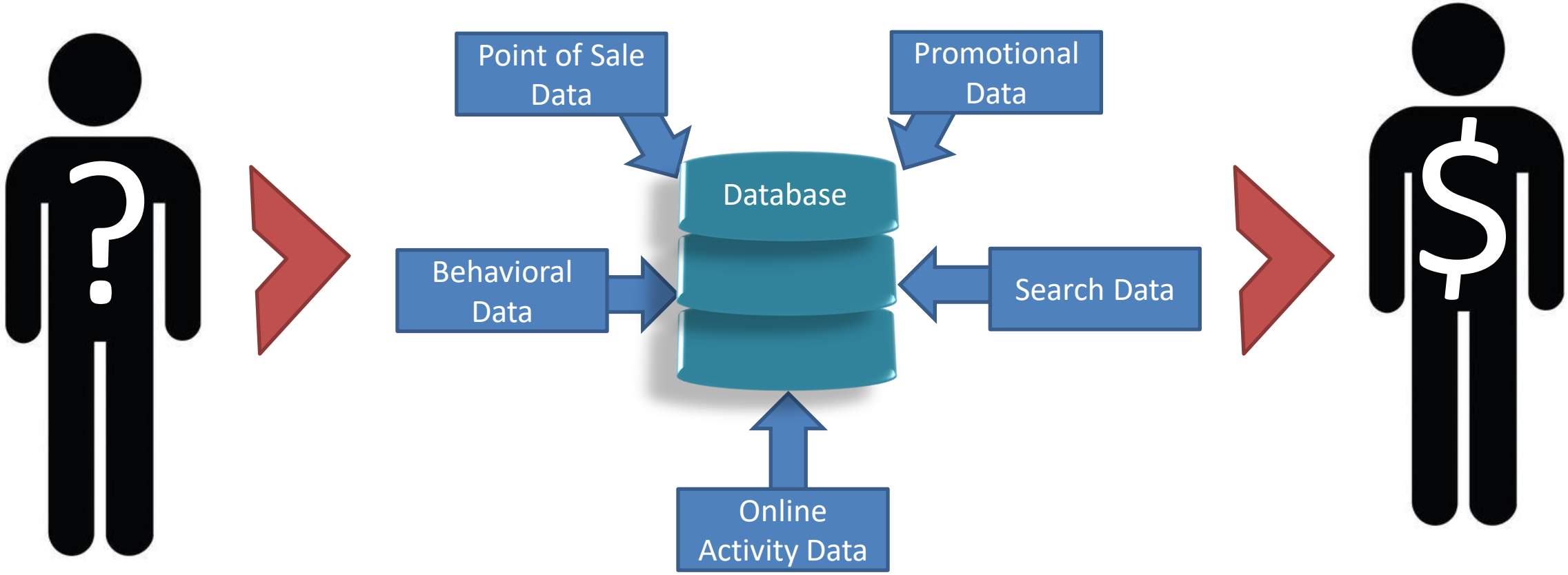


Preparation for Week 4: Predictive Analytics Applications: Predictive Lead Scoring

Target the lead who is most likely to convert and will be the highest value.
Algorithms can use explicit and implicit data from CRM.

Preparation for Week 4: Predictive Analytics Applications: Cross-Sell / Up-Sell

Problem: Customer acquisition costs are high



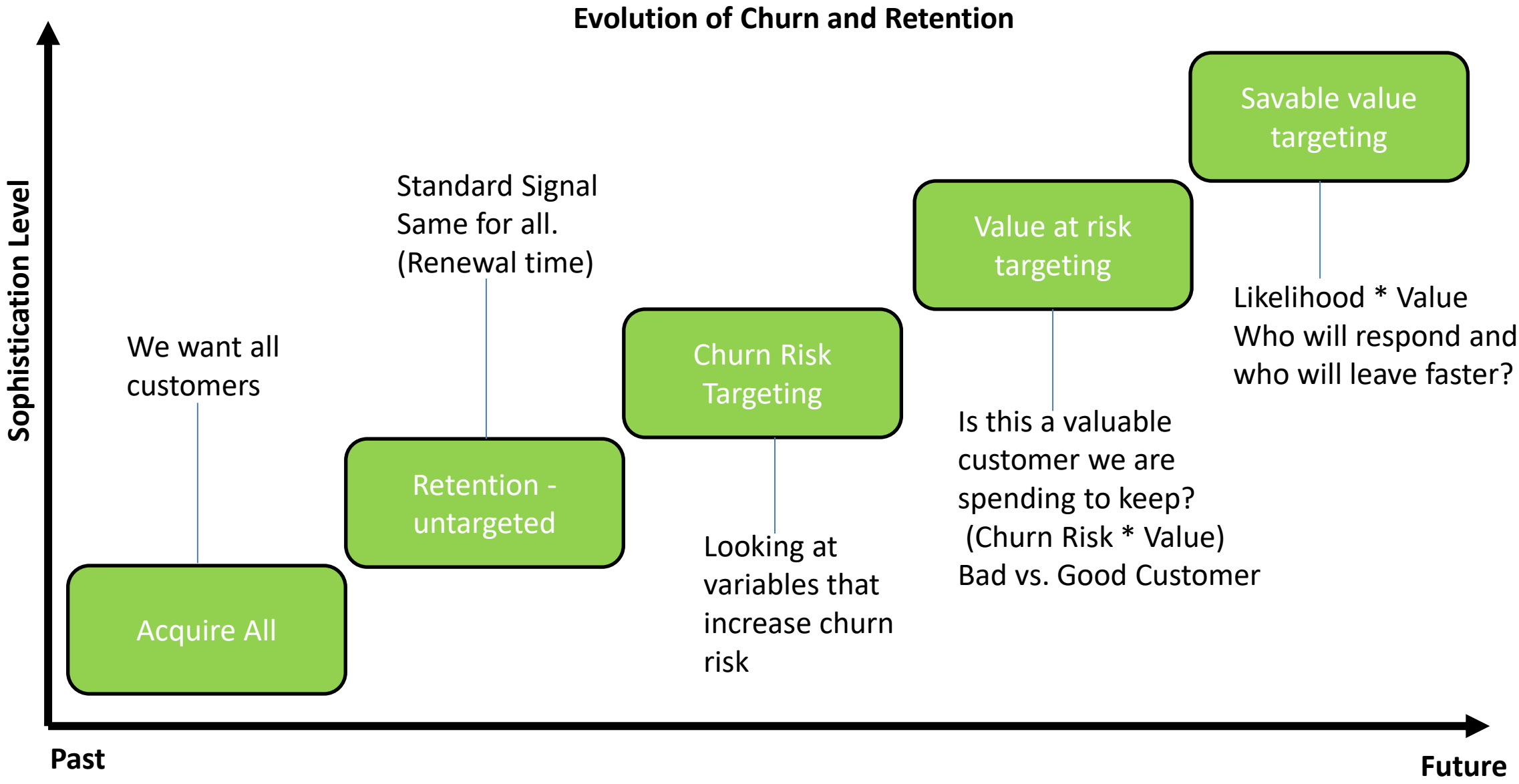
We don't know customer well until data is collected

Preparation for Week 4: Predictive Analytics Applications

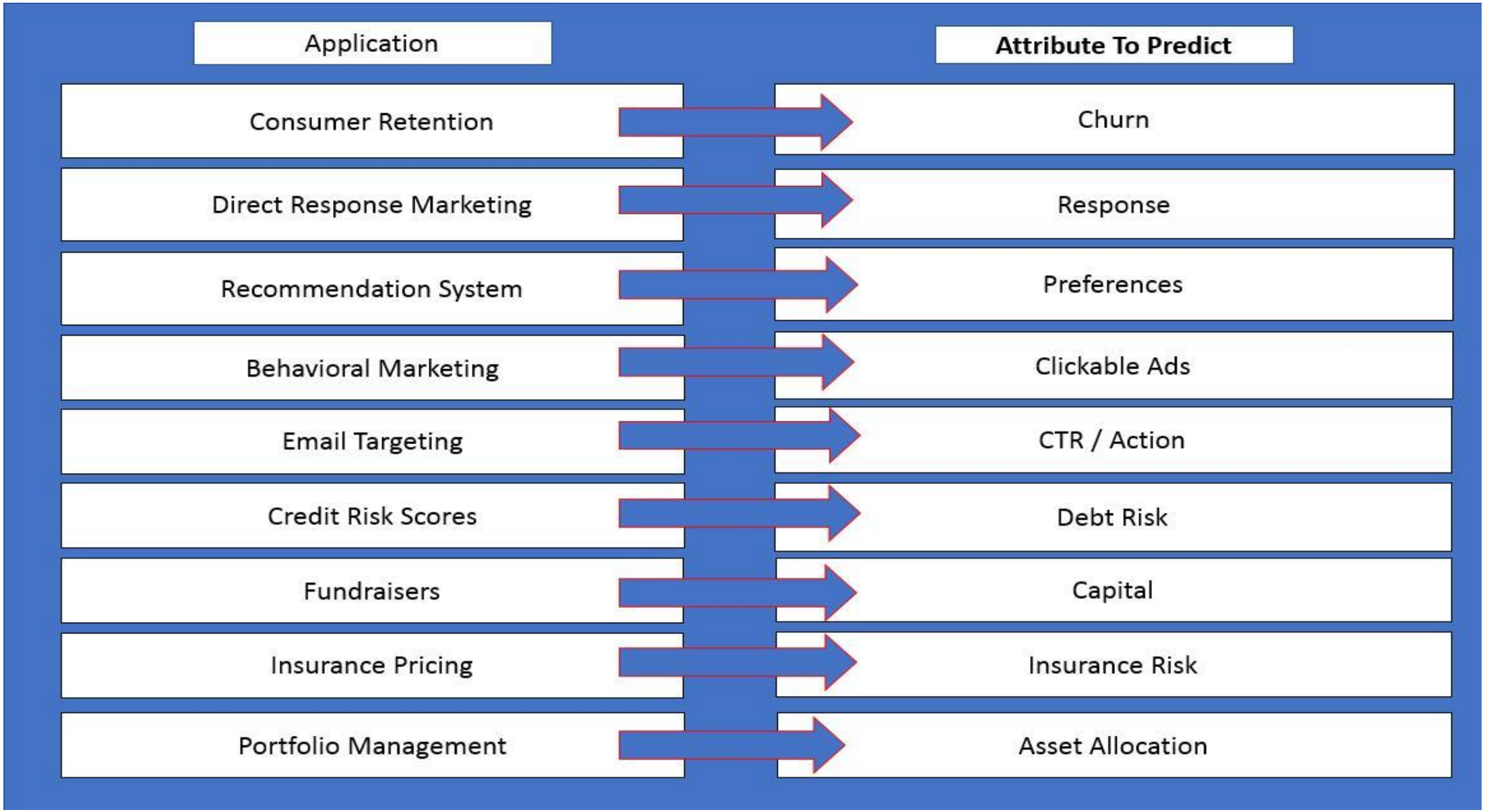
Attrition/Churn/Retention

Source: https://klaviyo.com

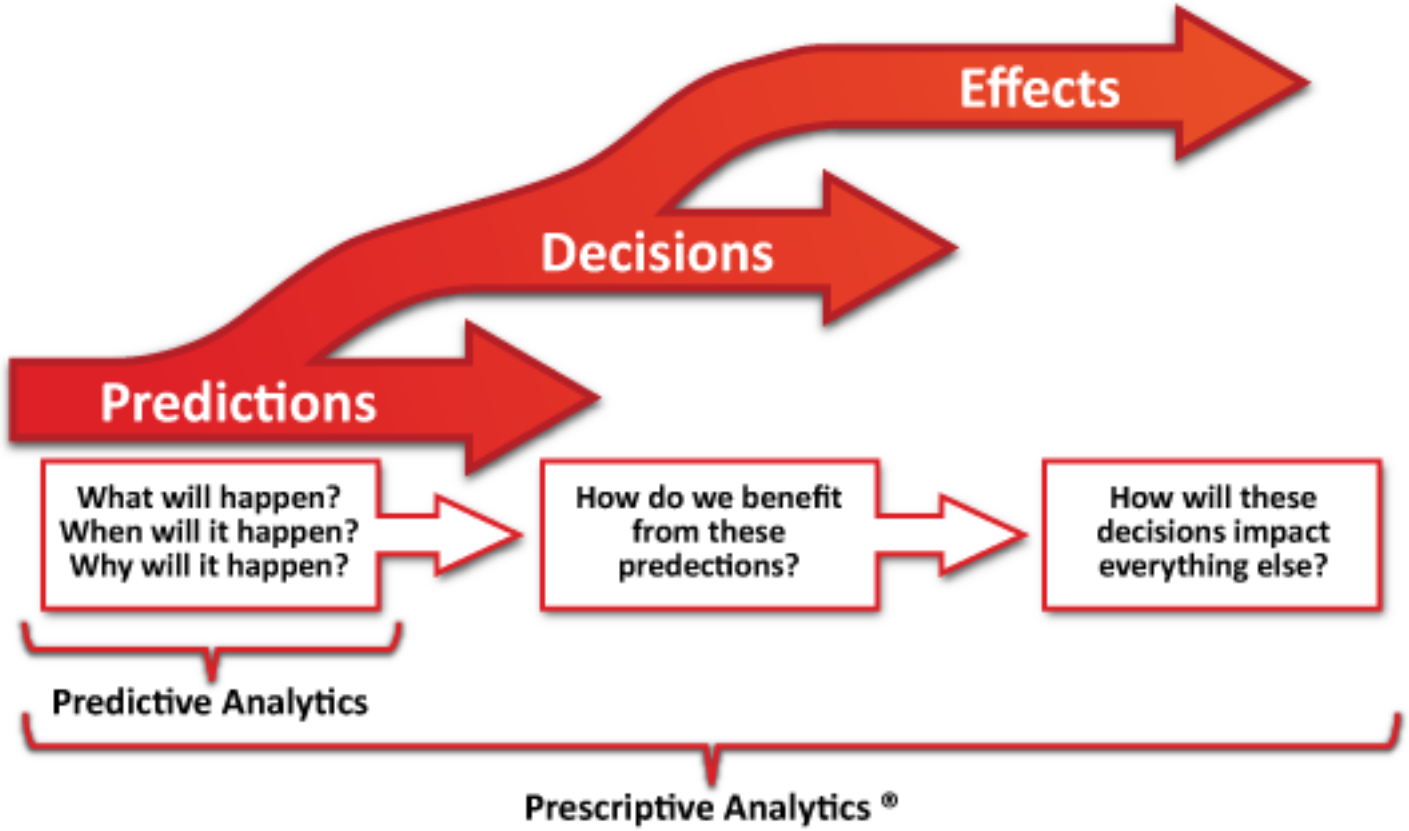
Preparation for Week 4: Predictive Analytics Applications – Customer Churn



Preparation for Week 4: Predictive Analytics Applications



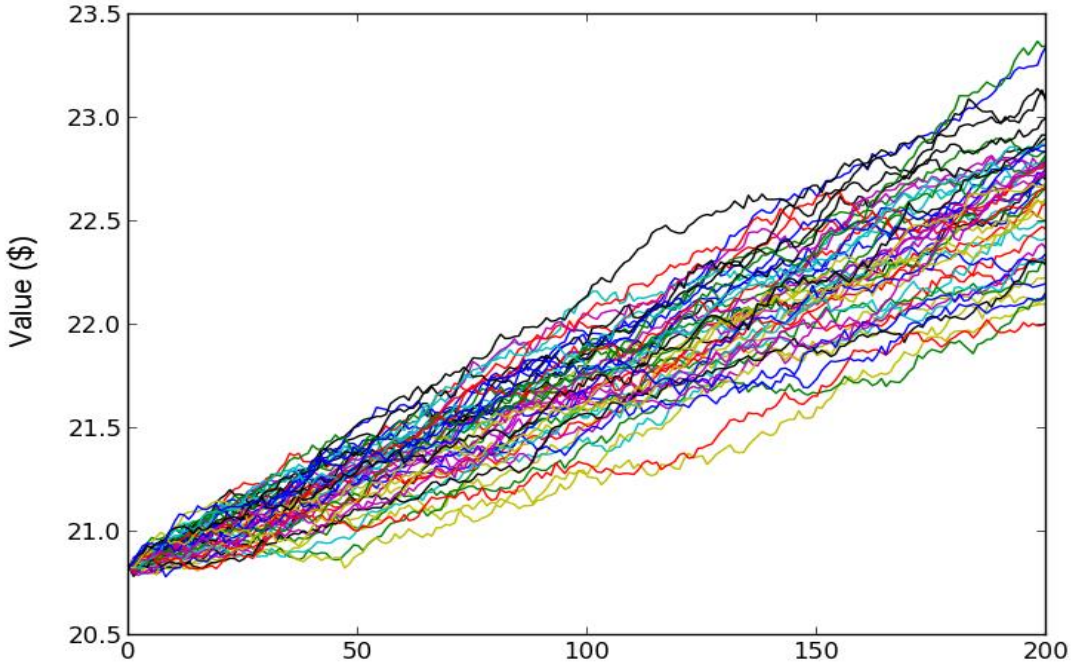
Preparation for Week 4: Predictive Analytics Applications – Risk Reduction

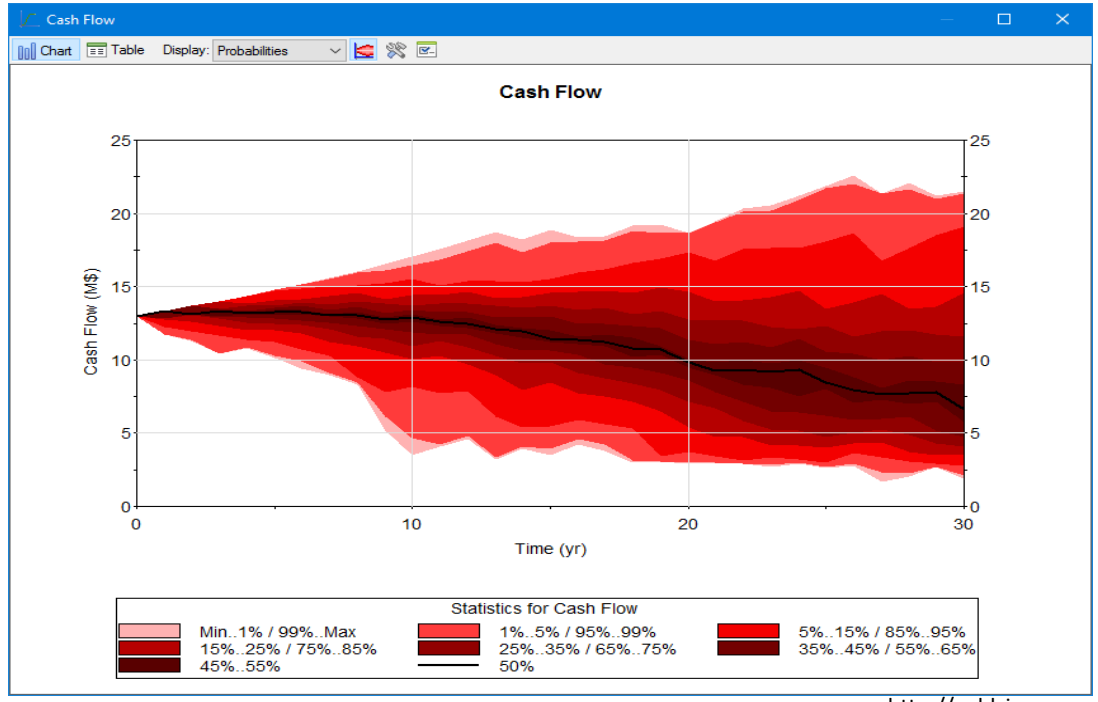


Source: https://commons.wikimedia.org/

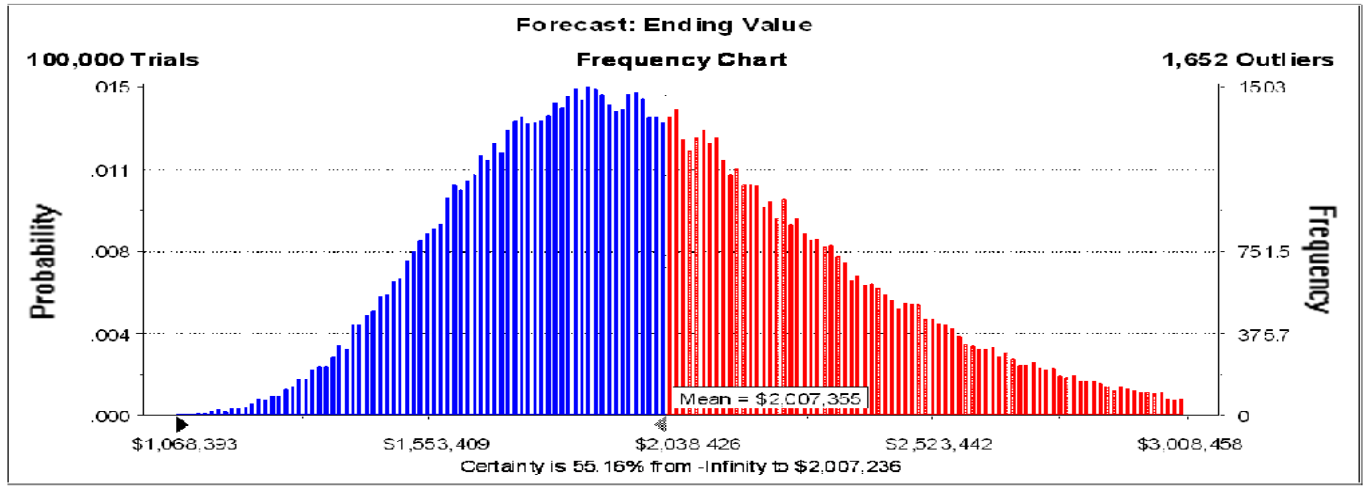
Preparation for Week 4: Predictive Analytics Applications – Financial Modeling

Simulated paths of the value of an asset using Monte Carlo



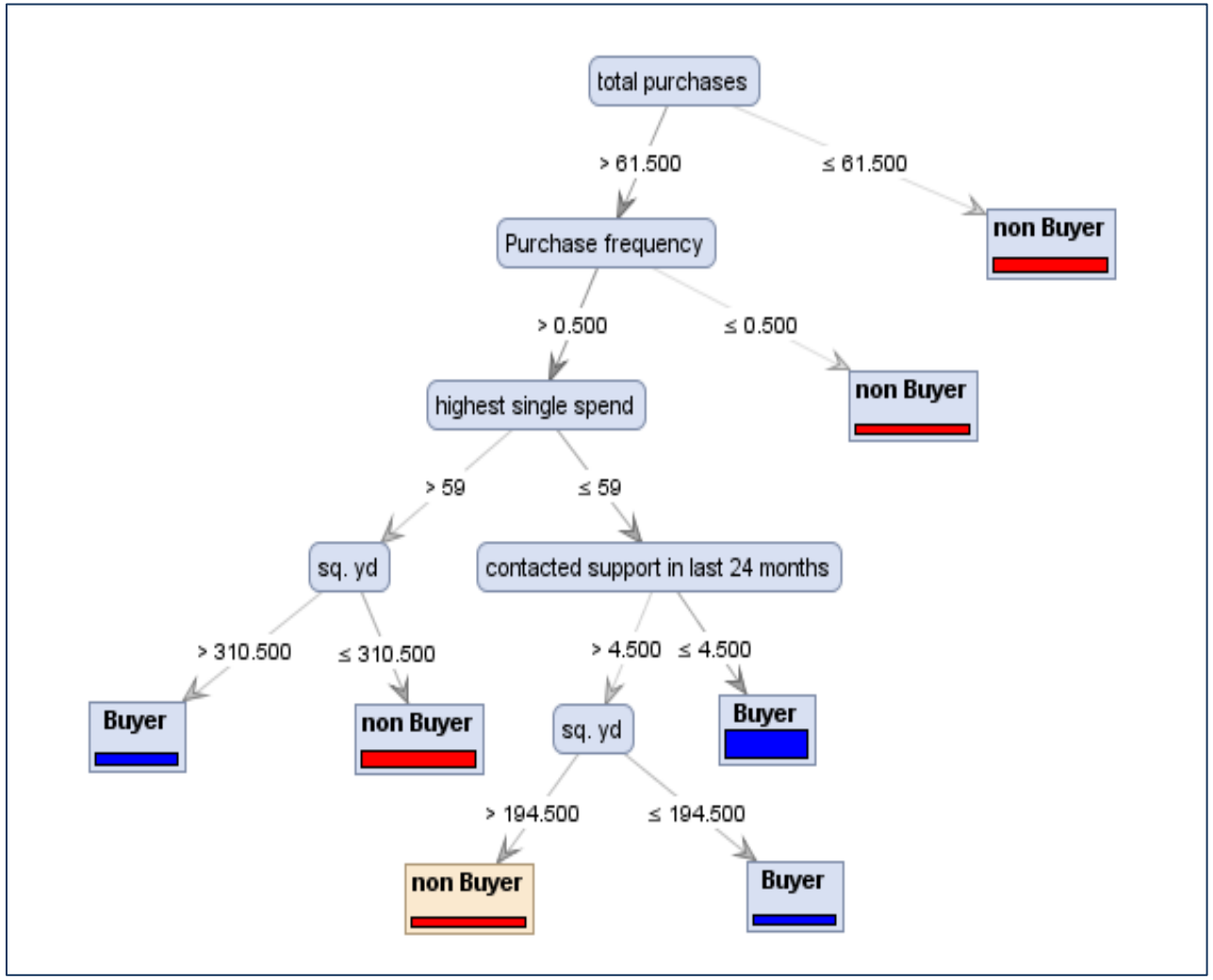


http://goldsim.com

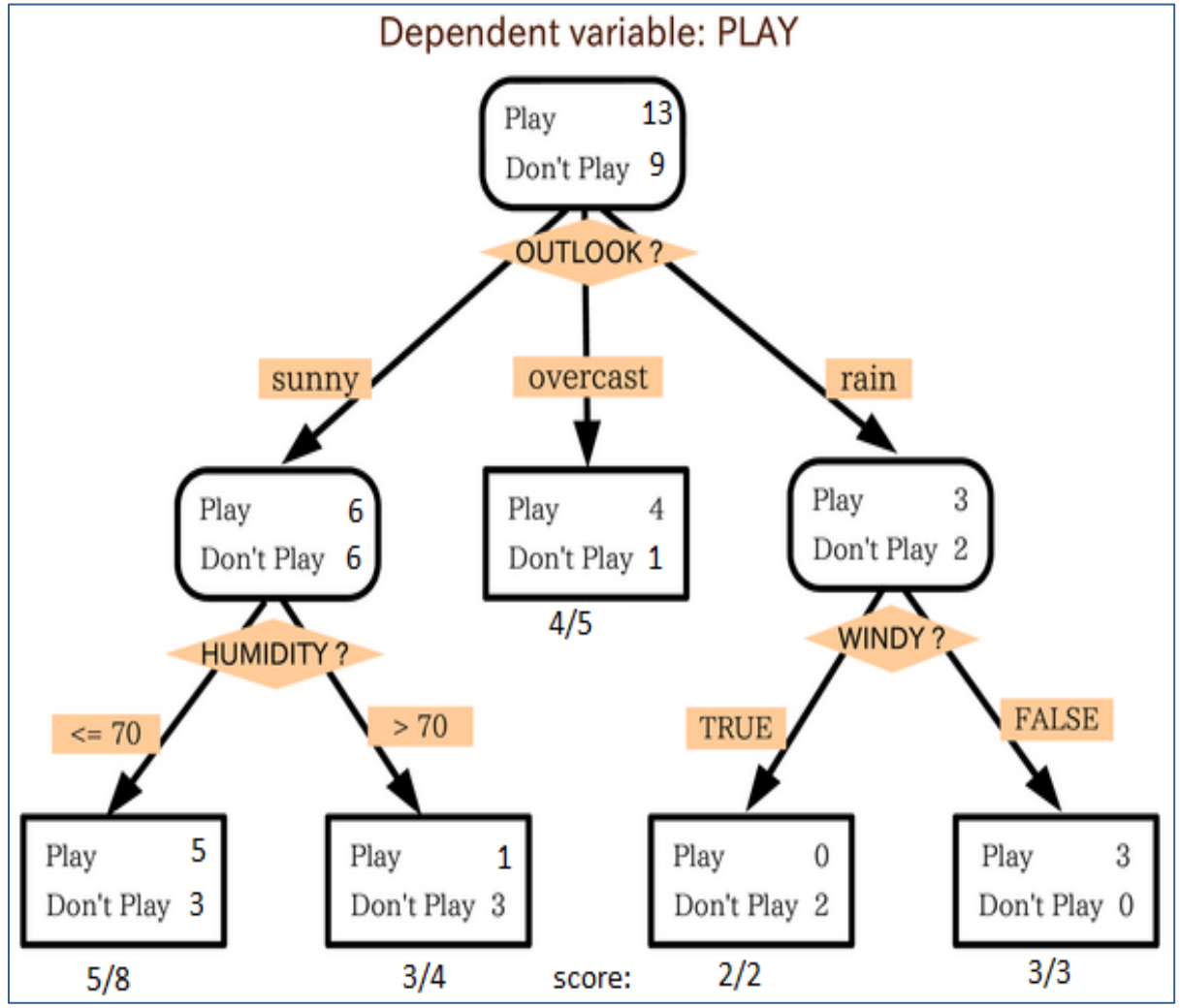


Source: Financialplanningbodyofknowledge.com

Preparation for Week 4: Predictive Analytics Applications – Risk Reduction

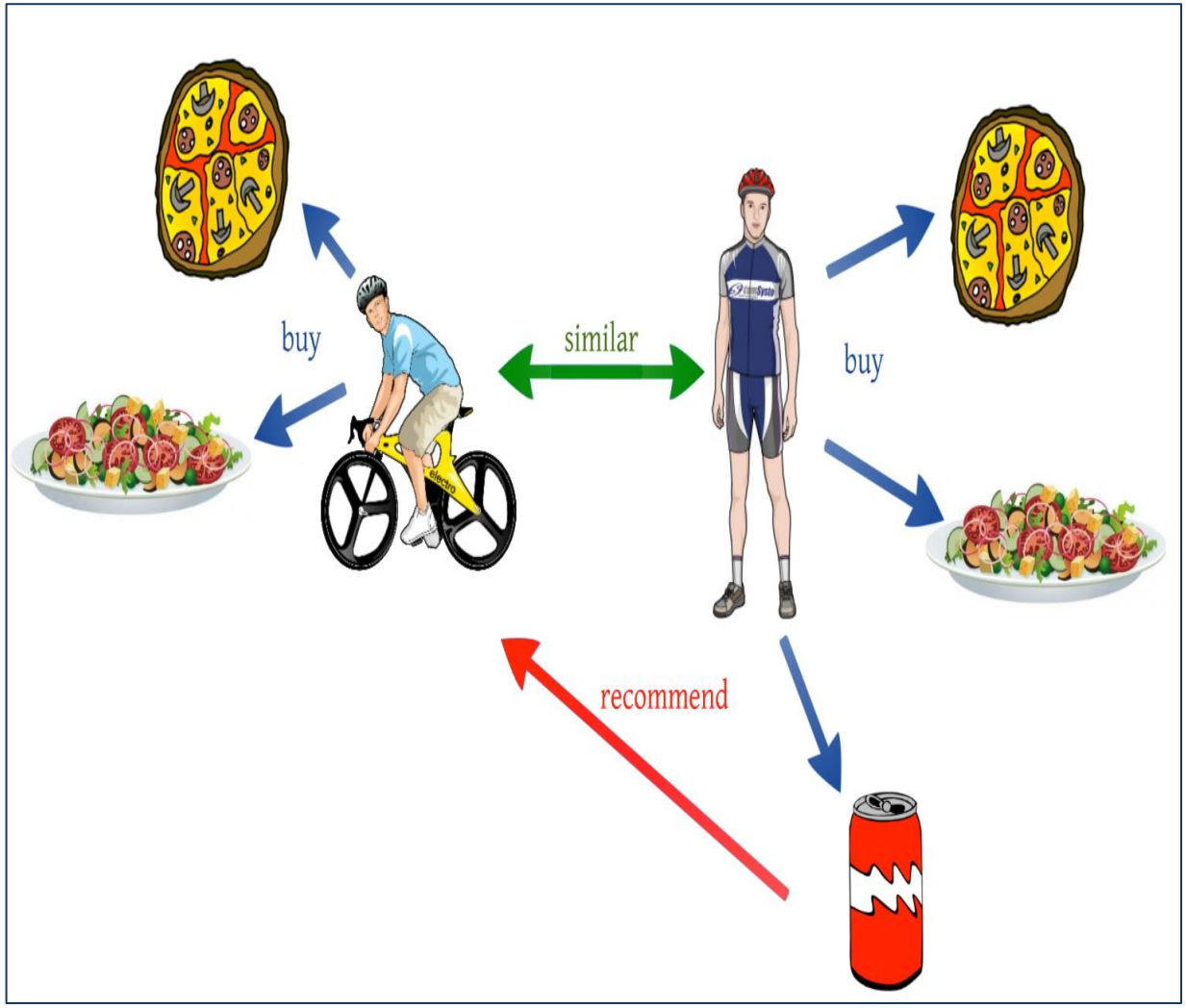


Source: http://simafore.com

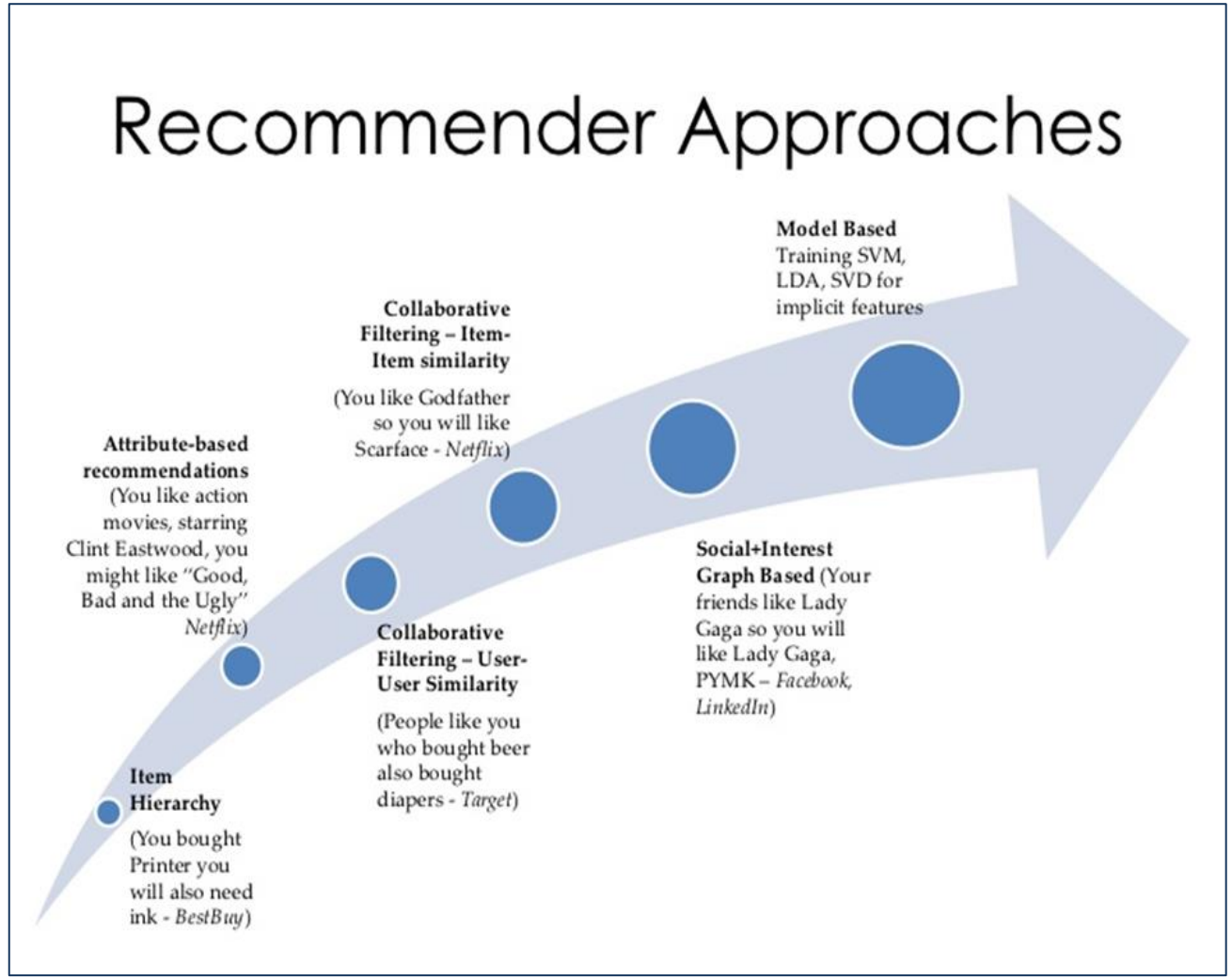


Source: http://stackexchange.com

Preparation for Week 4: Predictive Analytics Applications – Recommendations

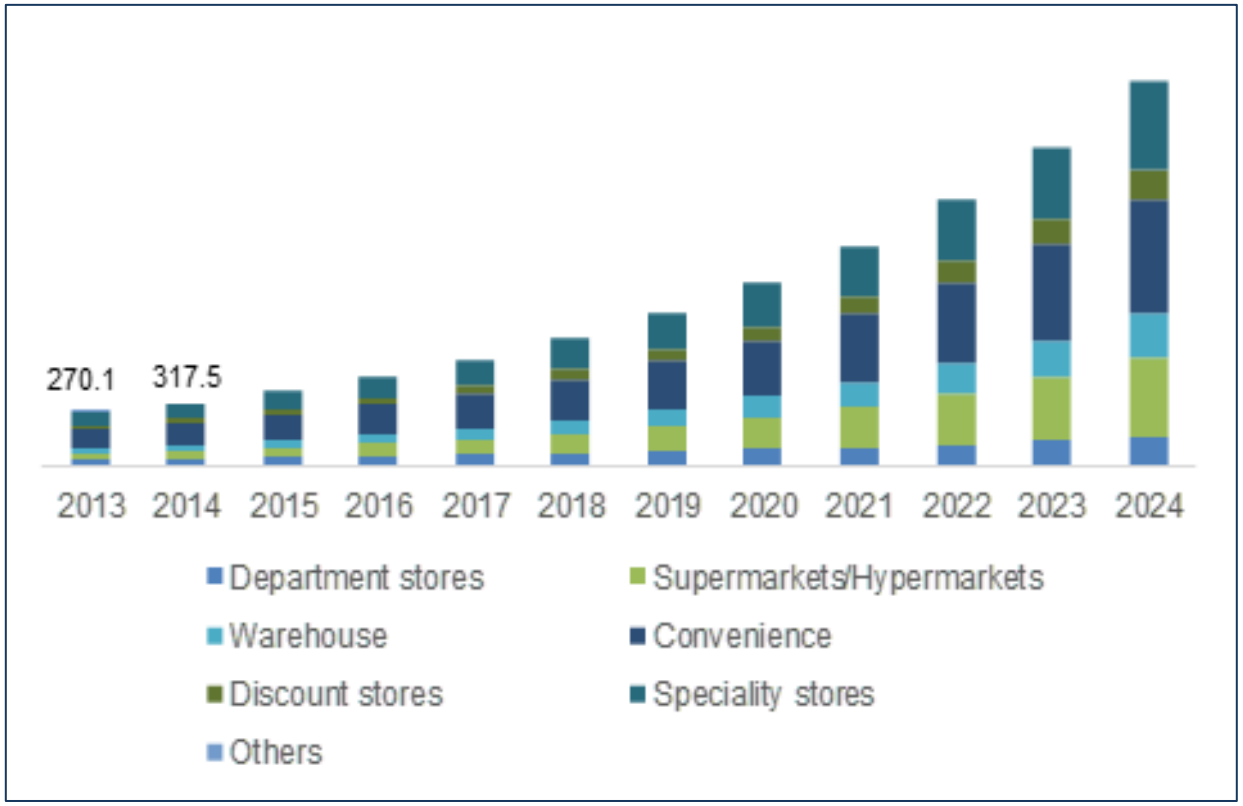


Source: http://medium.com

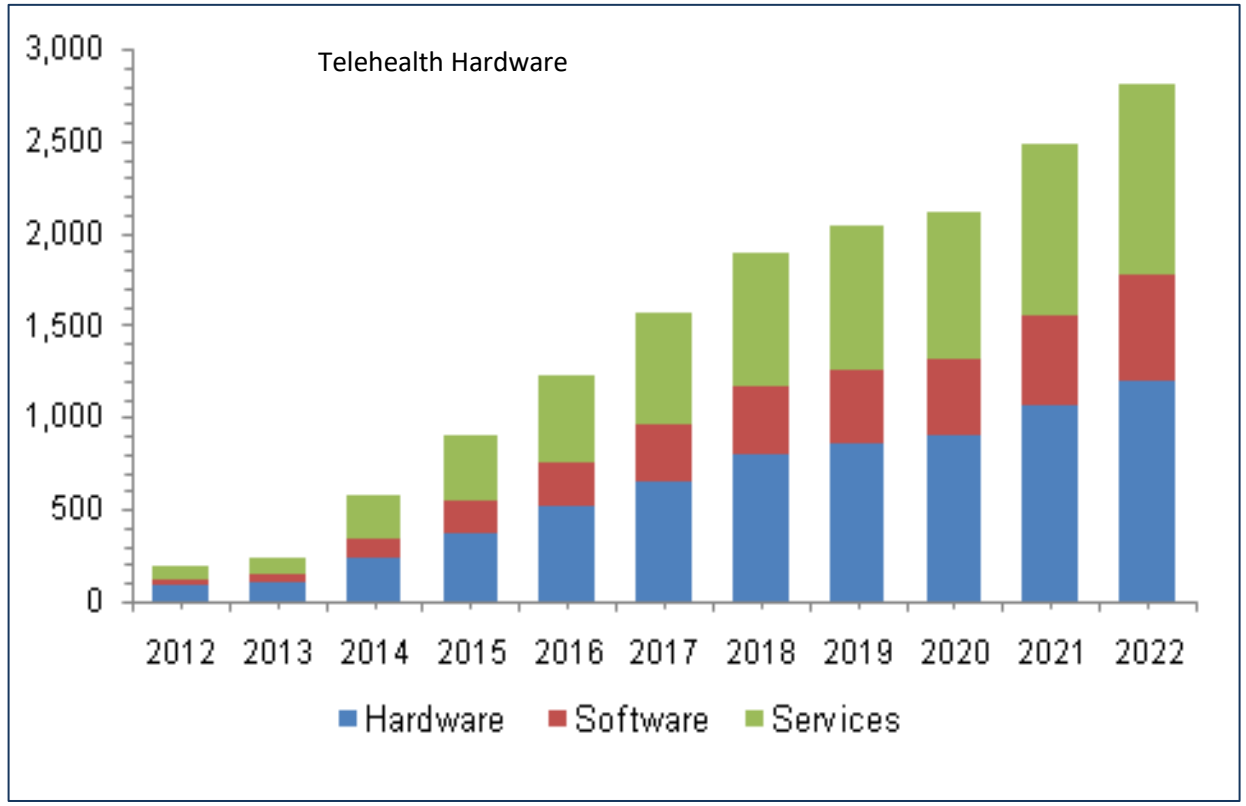


Source: http://Sflscientific.com

Preparation for Week 4: Predictive Analytics Applications – Market Analysis

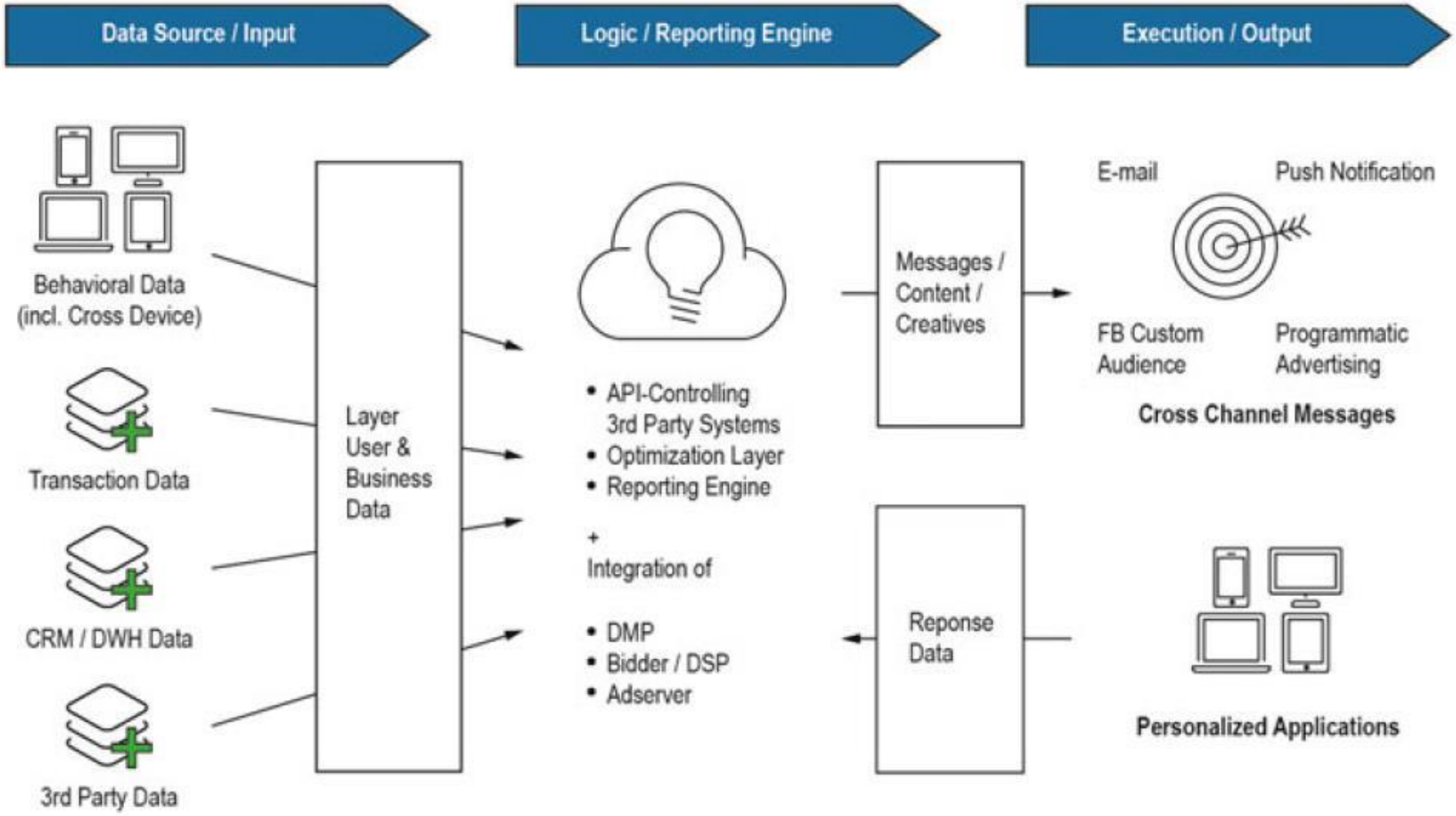


Source: http://www.gminsights.com



Source: http://chironhealth.com

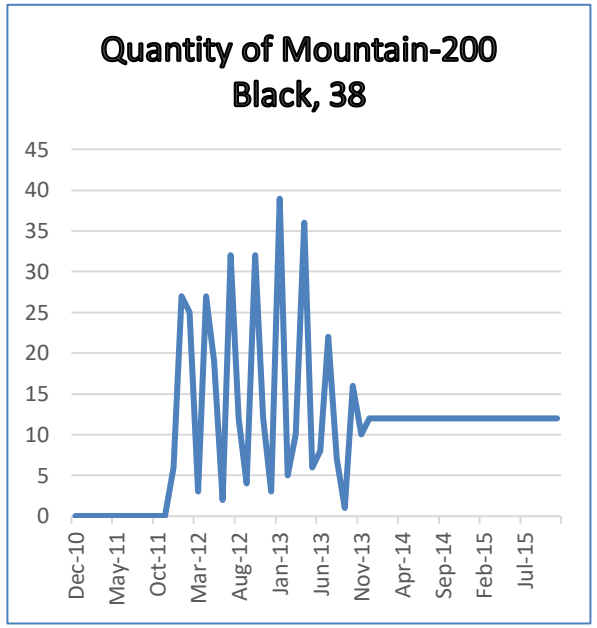
Preparation for Week 4: Predictive Analytics Applications – Digital Advertising



Source: Heinemann, F.(2016) Driving Performance with Programmatic CRM

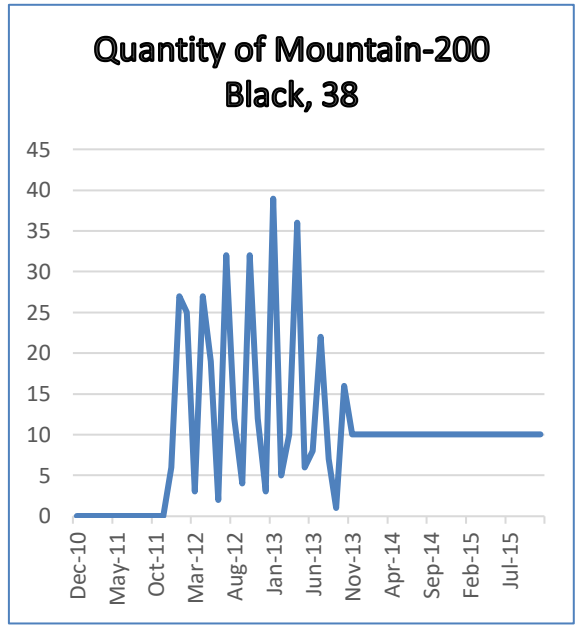
Preparation for Week 4: Time Series

Single Exponential Smoothing



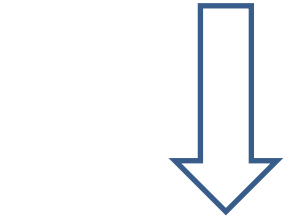
Mean Absolute Percentage Error (MAPE)	182.4256
Mean Absolute Deviation (MAD)	9.047372
Mean Square Error (MSE)	144.9178
Tracking Signal Error (TSE)	6.624276
Cumulative Forecast Error (CFE)	59.93229
Mean Forecast Error (MFE)	1.872884

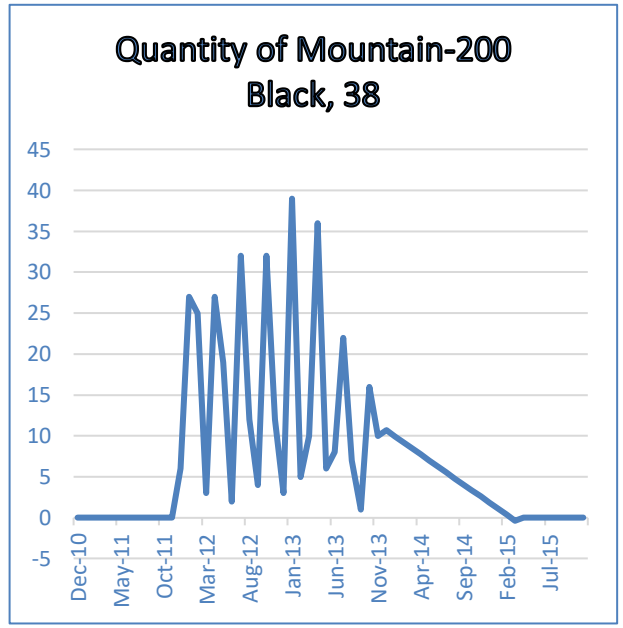
Moving Averages Smoothing



Mean Absolute Percentage Error (MAPE)	223.5788
Mean Absolute Deviation (MAD)	12.625
Mean Square Error (MSE)	290.5625
Tracking Signal Error (TSE)	0.792079
Cumulative Forecast Error (CFE)	10
Mean Forecast Error (MFE)	0.3125

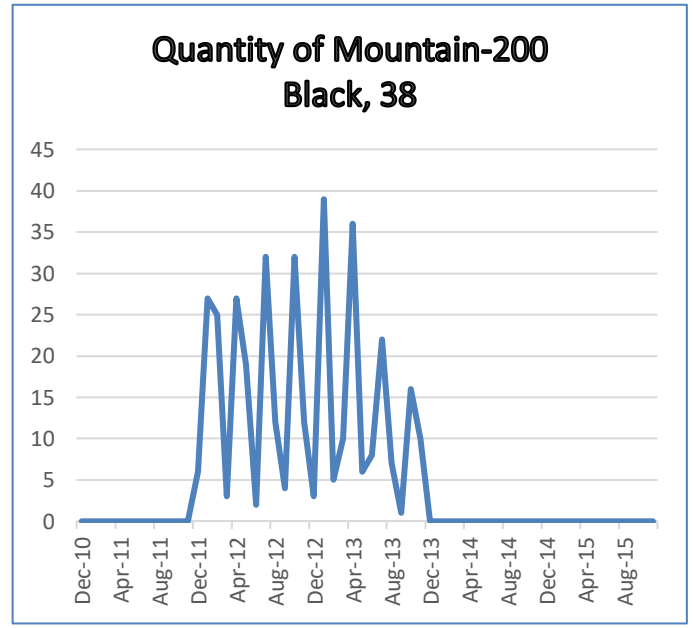
Double Exponential Smoothing





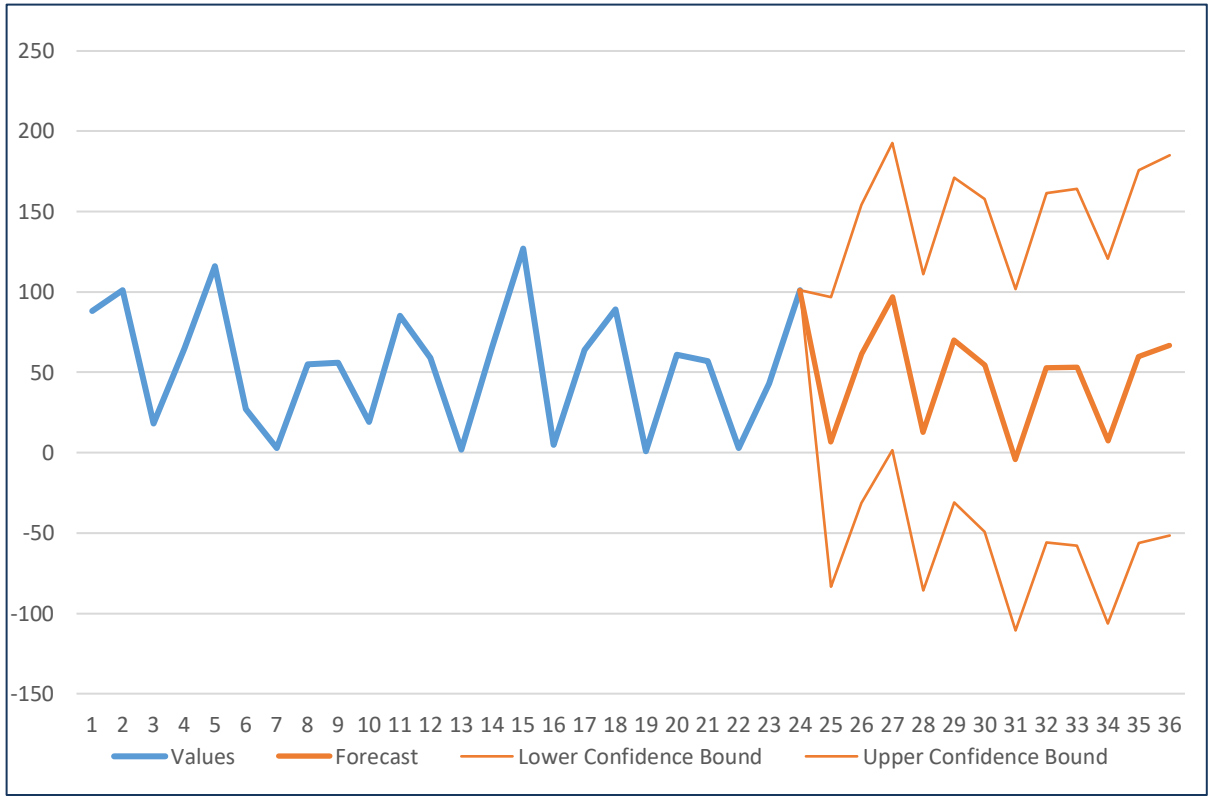
Mean Absolute Percentage Error (MAPE)	234.7263
Mean Absolute Deviation (MAD)	9.811871
Mean Square Error (MSE)	140.8039
Tracking Signal Error (TSE)	-6.78919
Cumulative Forecast Error (CFE)	-66.6146
Mean Forecast Error (MFE)	-2.01862

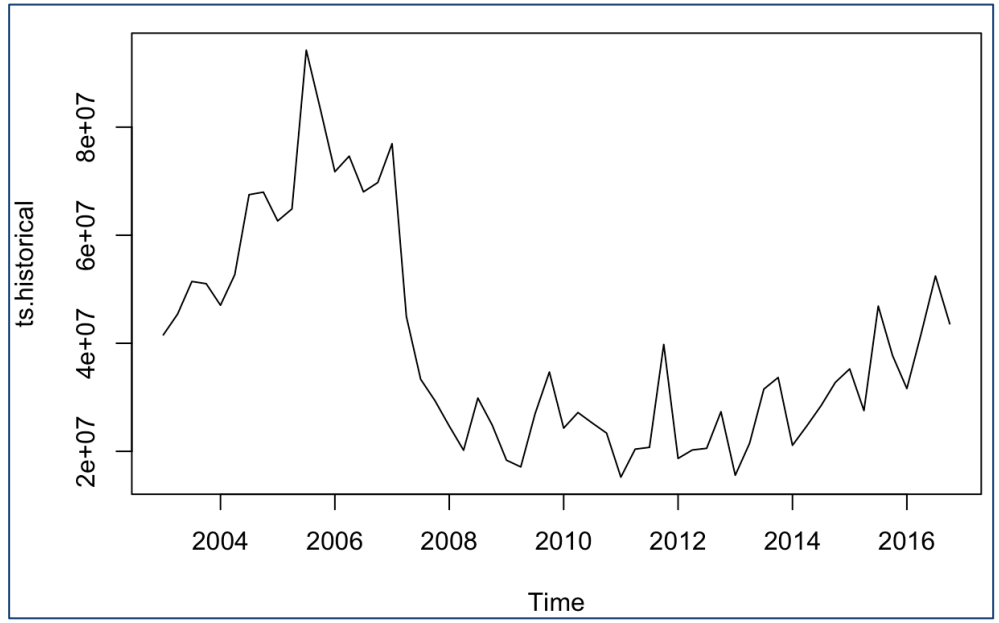
Triple Exponential Smoothing

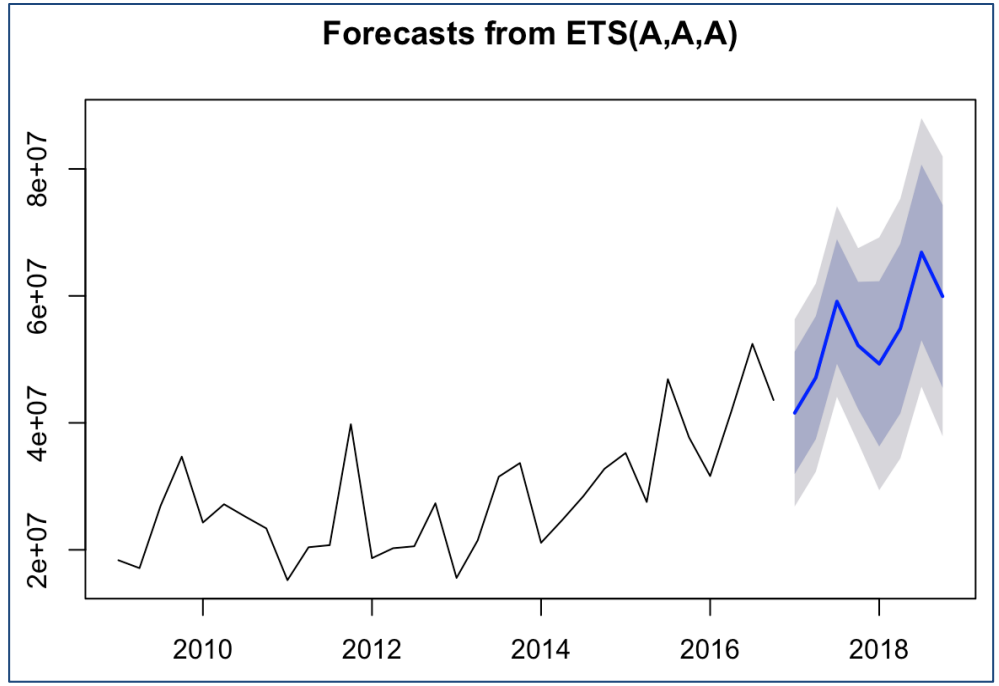


Mean Absolute Percentage Error (MAPE)	100
Mean Absolute Deviation (MAD)	11.0303
Mean Square Error (MSE)	266.2424
Tracking Signal Error (TSE)	33
Cumulative Forecast Error (CFE)	364
Mean Forecast Error (MFE)	11.0303

Preparation for Week 4: Tutorial For Predictive Analytics: Time Series Visualization







Preparation for Week 4: Predictive Analytics Applications – Multiple Regression

```
Residuals:
      Min        1Q    Median        3Q       Max
-20614597  -1849007         0   1740805  14013527

Coefficients:
                               Estimate      Std. Error t value          Pr(>|t|)
(Intercept)                55087544.17370  59902286.93867   0.920           0.36053
BuildingClassFinalRollC0    -511829.05153   5291132.55548  -0.097           0.92318
BuildingClassFinalRollC1    1972963.67852   5201342.63481   0.379           0.70546
BuildingClassFinalRollC2    -915817.86150   5416380.66169  -0.169           0.86616
BuildingClassFinalRollC3    1568124.41688   5343466.16755   0.293           0.76993
BuildingClassFinalRollC4     787887.78238   5025496.19446   0.157           0.87581
BuildingClassFinalRollC6   -5369646.61639   6332651.59816  -0.848           0.39901
BuildingClassFinalRollC7    1807141.72312   5225211.73749   0.346           0.73036
BuildingClassFinalRollD1    4888684.81144   6252896.69410   0.782           0.43662
BuildingClassFinalRollD3    6666868.66283   7723701.02615   0.863           0.39062
BuildingClassFinalRollD6  -21913591.13592   6607430.12443  -3.317           0.00137 **
BuildingClassFinalRollD7    2409822.25663   5797734.15671   0.416           0.67878
BuildingClassFinalRollD9    8022386.72043   6343883.57492   1.265           0.20969
BuildingClassFinalRollR9   10255415.14753   8060905.28646   1.272           0.20697
BuildingClassFinalRollRR   11899477.69424   7855328.16592   1.515           0.13376
ResidentialUnits            -248544.32987     98603.65269  -2.521           0.01370 *
CommercialUnits              276046.95689    646323.23727   0.427           0.67045
GrossSqFt                       794.17103        95.95051   8.277 0.00000000000232 ***
SaleDate                          0.03094         0.01027   3.012           0.00348 **
YearBuilt                       -50845.34076    31243.58536  -1.627           0.10759
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 4819000 on 80 degrees of freedom
Multiple R-squared:  0.8447,    Adjusted R-squared:  0.8079
F-statistic: 22.91 on 19 and 80 DF,  p-value: < 0.00000000000000022
```

Why we will use Adjusted R-squared

- R_2 will often increase if you add more predictors to the model
- A large R-squared may be an outcome of adding more predictor to the model
- We make use of the adjusted R-squared to account for the number of predictor in the model

How do I know which variables to add?

Generally, additional variables should have:

- Theoretical validity
- Explanatory power
- Avoid high correlation with another explanatory variable
- >.6 correlation is the standard cut-off.

2 C

Preparation for Week 4: Group Discussion Forum W4

Please select just one of the bullet points listed below:

- Discuss a use case for a regression model that can be applied to your professional field
- Discuss a use case for smoothing techniques that can be applied to your professional field
- Discuss the significance of the different types of regression models
- Discuss the significance of the different types of smoothing techniques
- Discuss topics of your choice from Lecture 7 and Lecture 8
- Discuss topics of your choice from Evans Chapters 8 and 9

Preparation for Week 4: Quiz 4: Q & A

Quiz 4 consists of 10 multiple choice and true-false questions.

Quiz 4 will cover information from the
- Online lecture notes (Lecture 7 and Lecture 8),
- Online discussion Week 4.

The quiz is open-book.

Format
- You will have 30 minutes to complete the quiz. There is a clock in the upper right corner of the screen keeping time for the exam.
- You can take the quiz only once.
- Each question will be delivered one at a time.
- You can revisit the questions and change your answers as many times as you want before submitting the exam.

2 C

Preparation for Week 4: Module 4 Exercise: Q & A

Questions are located at the end of Chapter 8

Section: Problems and Exercises

Group 1 - Roman	
Due Date	Monday, June 5, 2023
Student	**Chapter 8**
Yurui Chen	7
Victor Brice Fedjo Yemele	8
Jonathan Garrison	9
Payton Hatcher	10
Faria Hossain	2
Jiamin Li	3
Kaitlynn Nguyen	4
Miranda Petrillo	5
Haoqiang Qi	6
Nesteshia Riddell-Dell	20
Cassandra Simoneau	21
Yang Yang	22
Samuka Yekeh	23

Group 2 - Vivek	
Due Date	Monday, June 5, 2023
Student	**Chapter 8**
Audrey Chan	7
Flori-Ann DeLa Cruz	8
Marcel Fernandes Silva	9
Avirul Islam	10
Jiarui Lin	2
Joyce Machau	3
Nicole Matarazzo	4
Timothy Olakunle	5
Olu Olayeye	6
Sandhya Ramani	20
Sri Amruta Sripada	21
Samuel Stevens	22
Jack Swartz	23

Business Running Case–Targeted Outcomes:

BOSTON UNIVERSITY

W-2	**Term Project A1**: Data Import and Visualization
	Term Project A2. Data Manipulation In R
W-3	**Term Project A3**. Descriptive Analytics Techniques In R
W-4	**Term Project A4**. Predictive Analytics Techniques In R
W-6	**Term Project A5**. Prescriptive Analytics Techniques in R and Excel
W-7	**Term Project A6**. **Term Project Final Presentation** Model Deployment

V-Lab Instructions

All course tools and SQL Server can be accessed from the V-Lab, which is included in tuition.

Assignment 4 Requires access to R Studio and SQL Server

Assignment 4: Predictive Analytics

Assignment 4 Objective: Prepare a managerial report, starting with an executive summary; expected length up to 4 pages APA format, excluding cover page, table of content, and appendixes.

1. Perform time series analysis on the total dollar amount of residential real estate sales on your neighborhood. Use sales beginning in the year 2009 to develop your model. Develop a forecast for the next 8 quarters of sales.
2. Use a multiple regression model to come up with another forecast for the next 8 quarters of sales. Include time and seasonality. Use sales beginning in the year 2009 to develop your model.
3. Use a multiple regression model to determine the sale of a given residential property in your neighborhood. Include:
    a. Sale Date
    b. Year built
    c. Building type (categorical)
    d. Gross Square Feet
    e. Number of Units
4. According to your model from (3), what are the most and least useful predictors of the amount of a sale?
5. Are there any redundant independent variables in your model from (3)? How can you tell?
6. According to your model from (3), which properties were the biggest bargains and which were the most overpriced? How might you account for these disparities?
7. Write 3-4 pages summarizing your findings with a focus on the output, interpretation of the output, and what the insights mean for our decision-making process

[Hint: your model may be more accurate if you discard data from prior to the housing market crash]

LOCATION OF THE INFORMATION AND SOFTWARE

Course Website

Page on Blackboard: "Assignments">>Assignment 4

Tutorials Section

Link
Tutorial – Assignment 4

Q & A