Week 6: Forecasting (Predictive Analytics) in Supply Chains (SCs)

Presented by:

Dr. Mehdi Rajabi Asadabadi

Predictive Analytics

Predictive analytics focus on **using historical data** to identify patterns enabling the **prediction of the future**.

The identified pattern or trend from historical data is represented by a **mathematical model**.

This model can then be used to predict future events based on the previous data and the new data.

Predictive Analytics Models

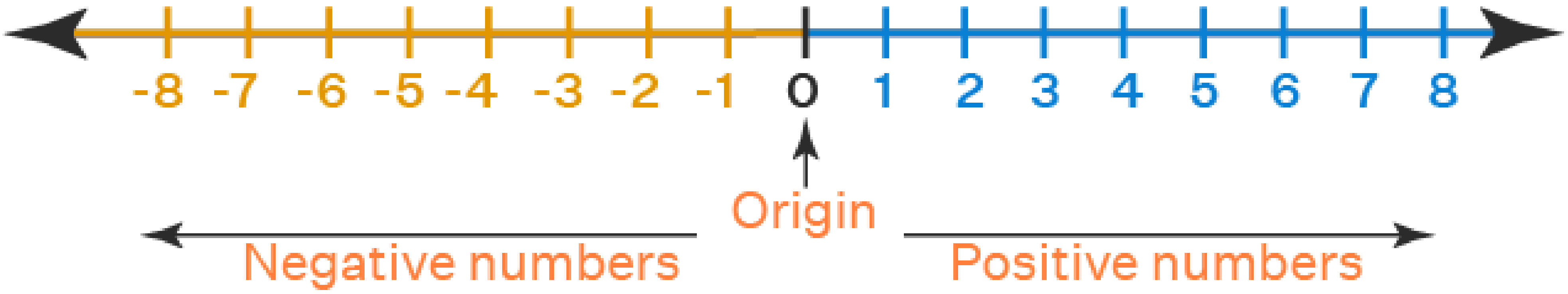
Classification:

Classification is the process of creating a set of classes for data, based on the existing data.

Predictive Analytics Models

Classification:

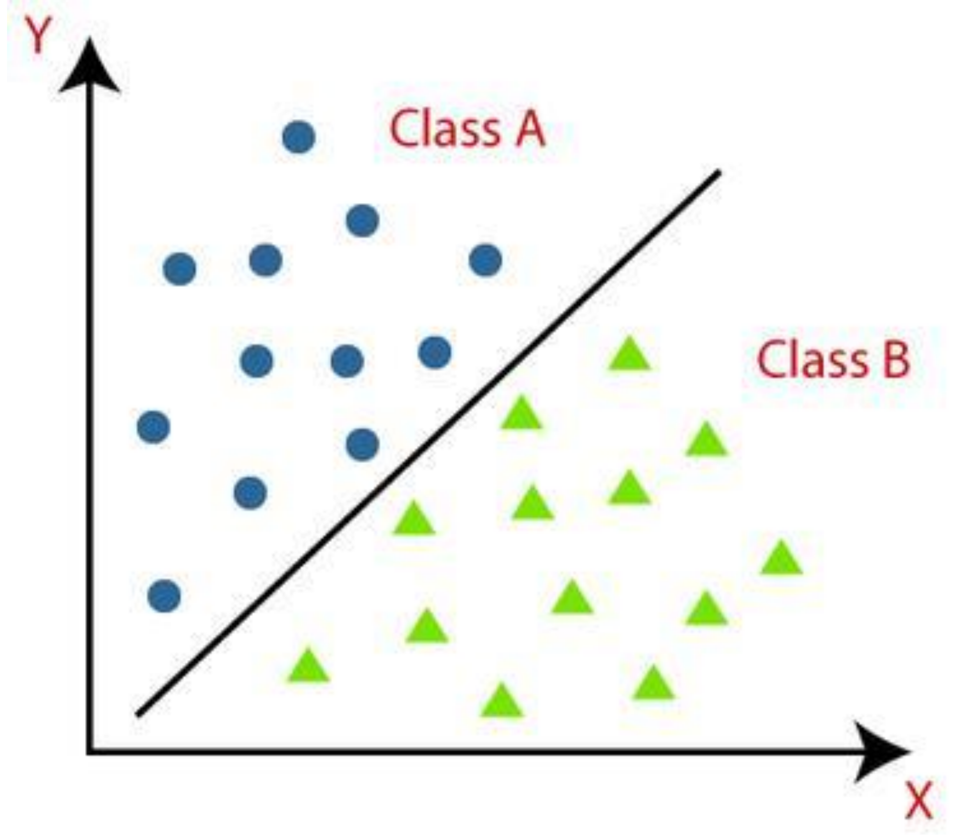
Classification is the process of creating a set of classes for data, based on the existing data.



Predictive Analytics Models

Classification:

Classification is the process of creating a set of classes for data, based on the existing data.

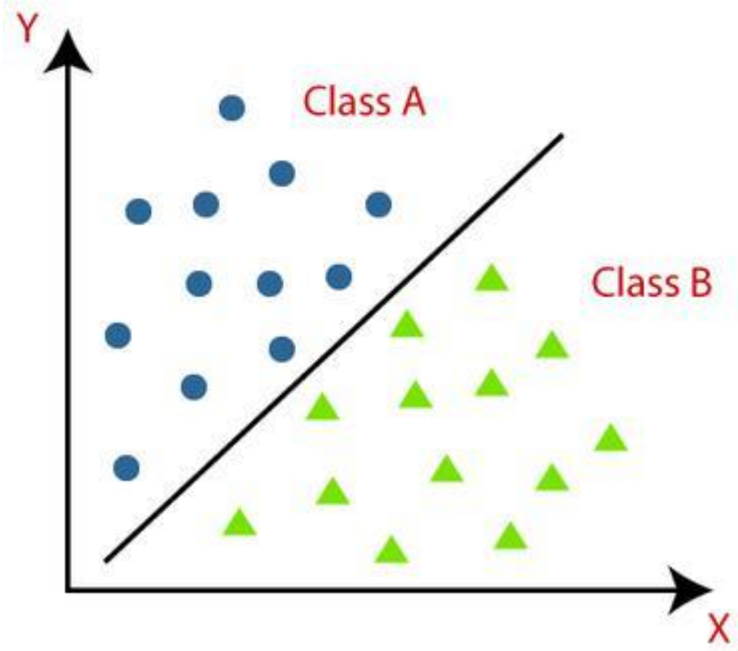


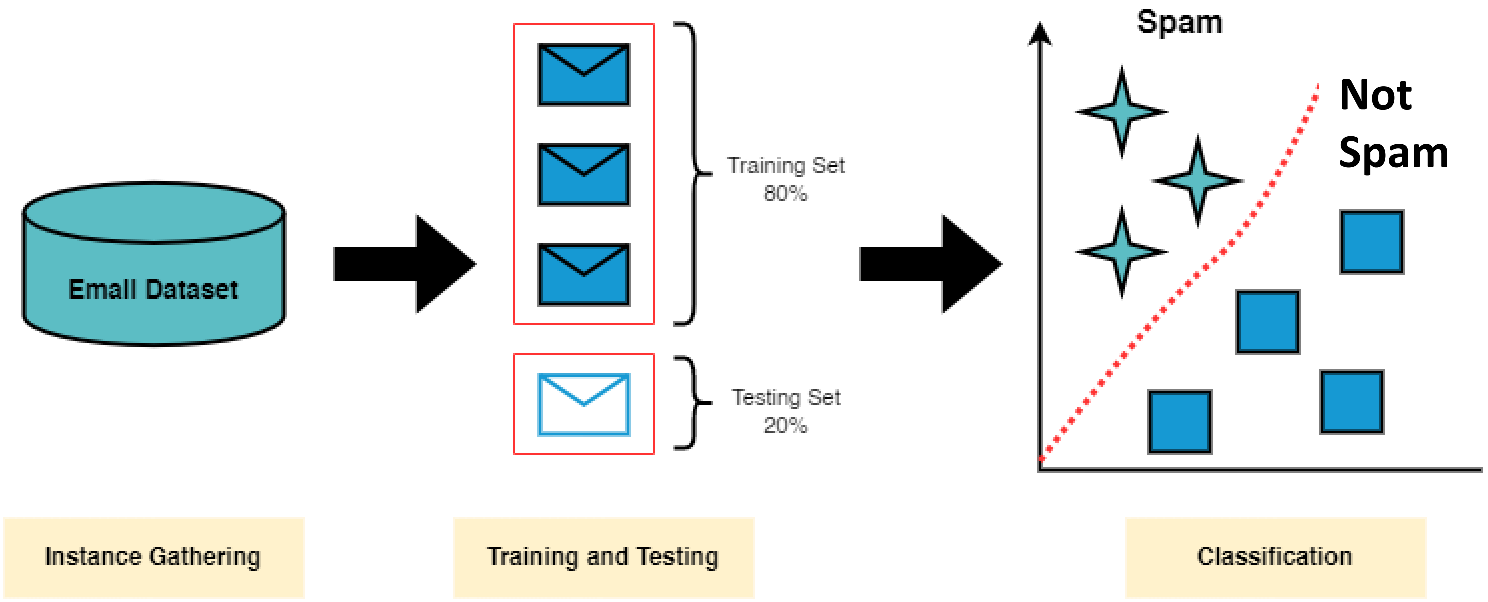
Predictive Analytics Models

Classification:

Classification is the process of creating a set of classes for data based on the existing data.

Binary classification:



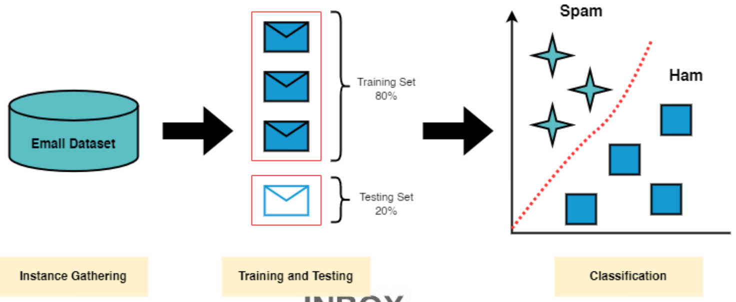


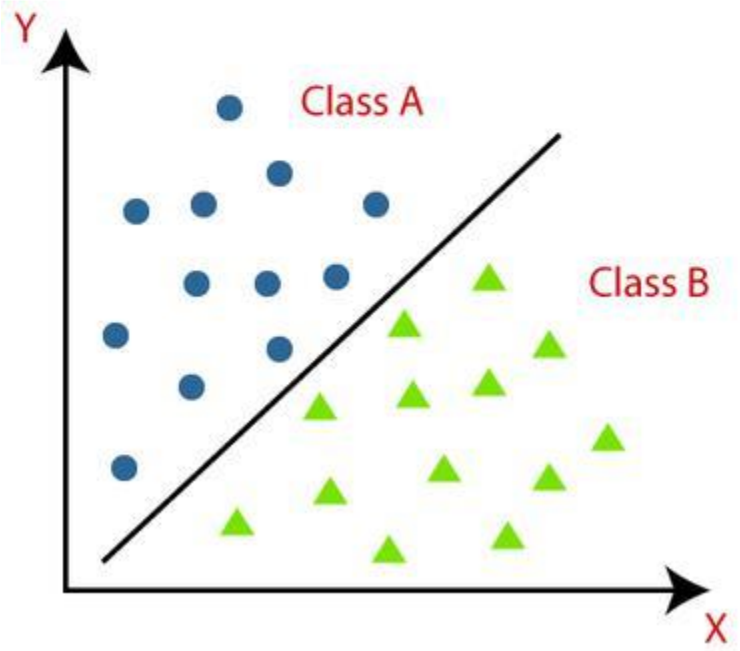
Predictive Analytics Models

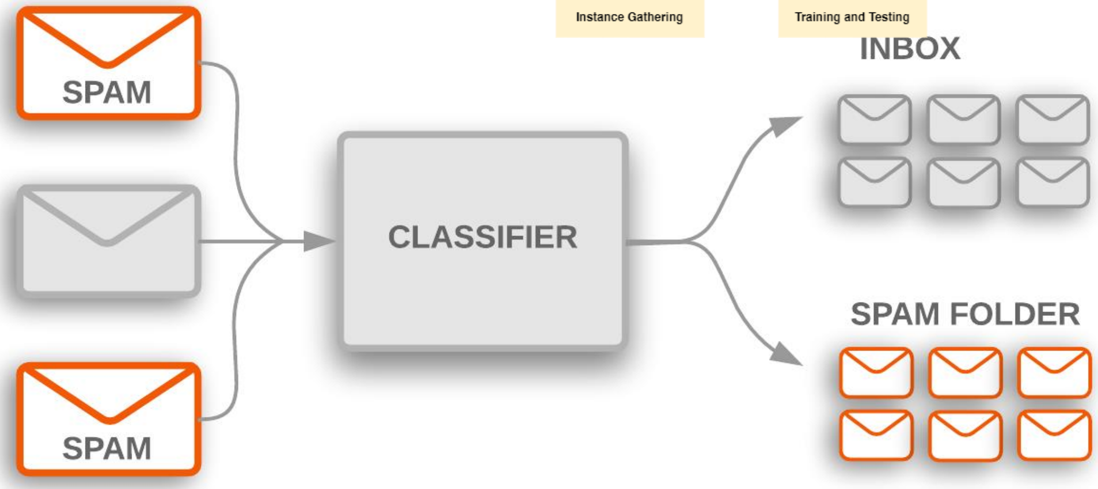
Classification:

Classification is the process of creating a set of classes for data based on the existing data.

Binary classification:







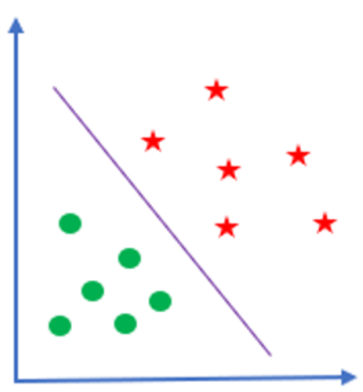
Predictive Analytics Models

Classification:

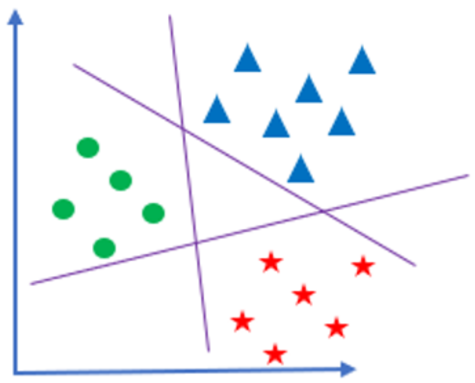
Classification is the process of creating a set of classes for data based on the existing data.

Multiclass classification:

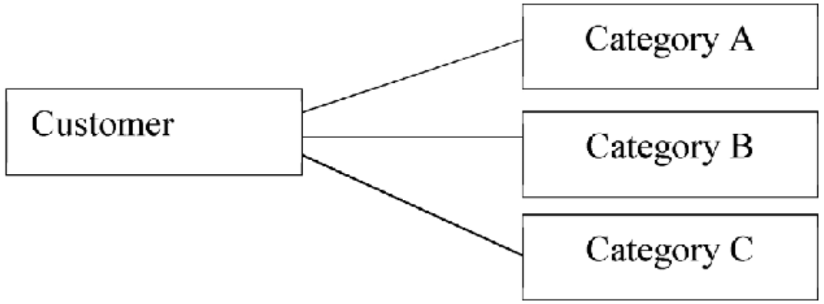
Binary classification

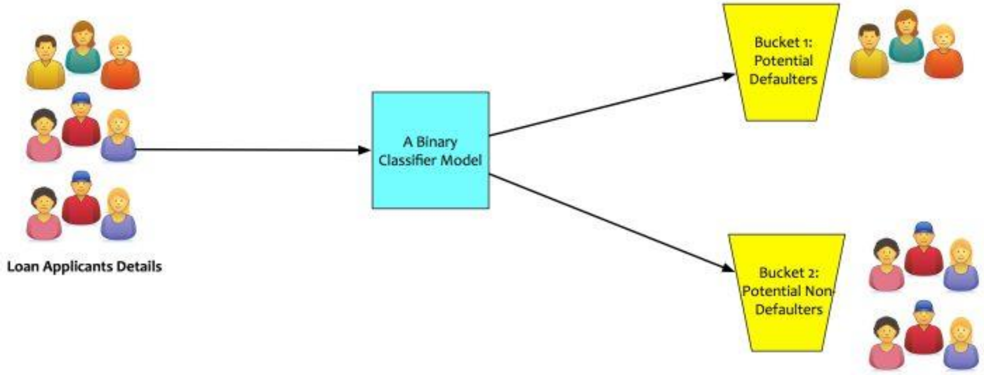


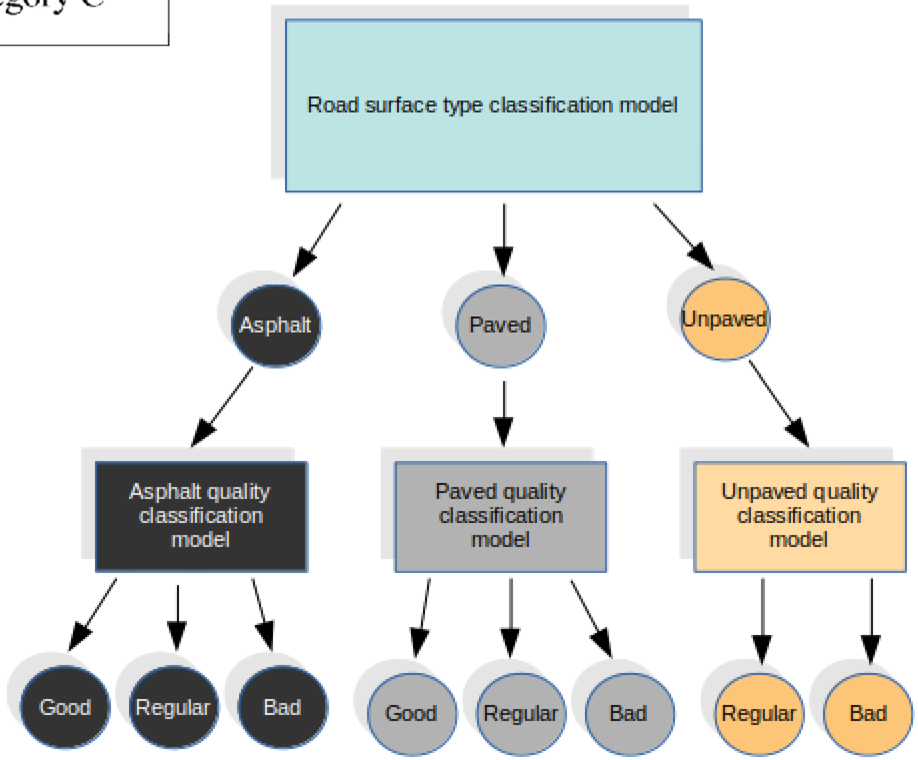
Multi-class classification

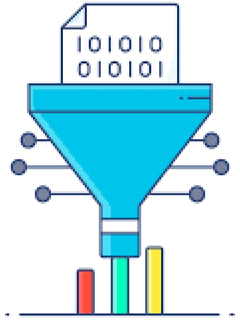


LIKELIHOOD ↓ \ SEVERITY →	1	2	3
1	LOW - 1 -	LOW - 2 -	MEDIUM - 3 -
2	LOW - 2 -	MEDIUM - 4 -	HIGH - 6 -
3	MEDIUM - 3 -	HIGH - 6 -	HIGH - 9 -

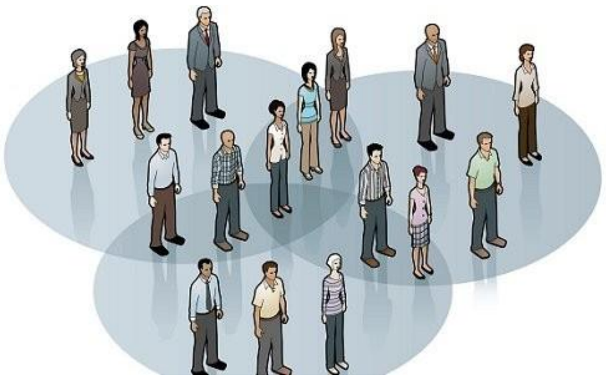








Big Data Filtering



Predictive Analytics Models

Classification:

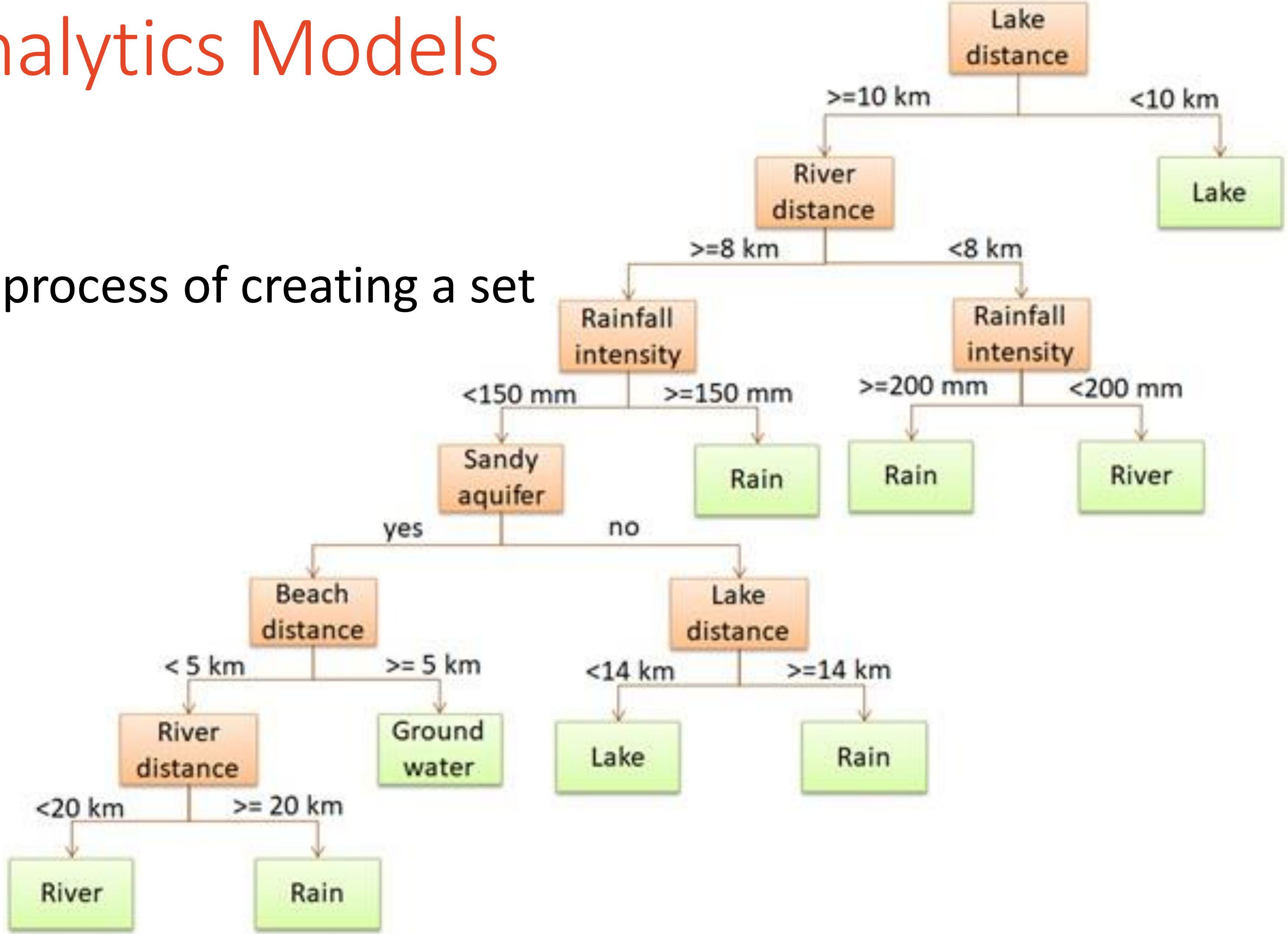
Classification is the process of creating a set of classes for data based on the existing data.

Predictive Analytics Models

Classification:

Classification is the process of creating a set of classes for data

Decision trees

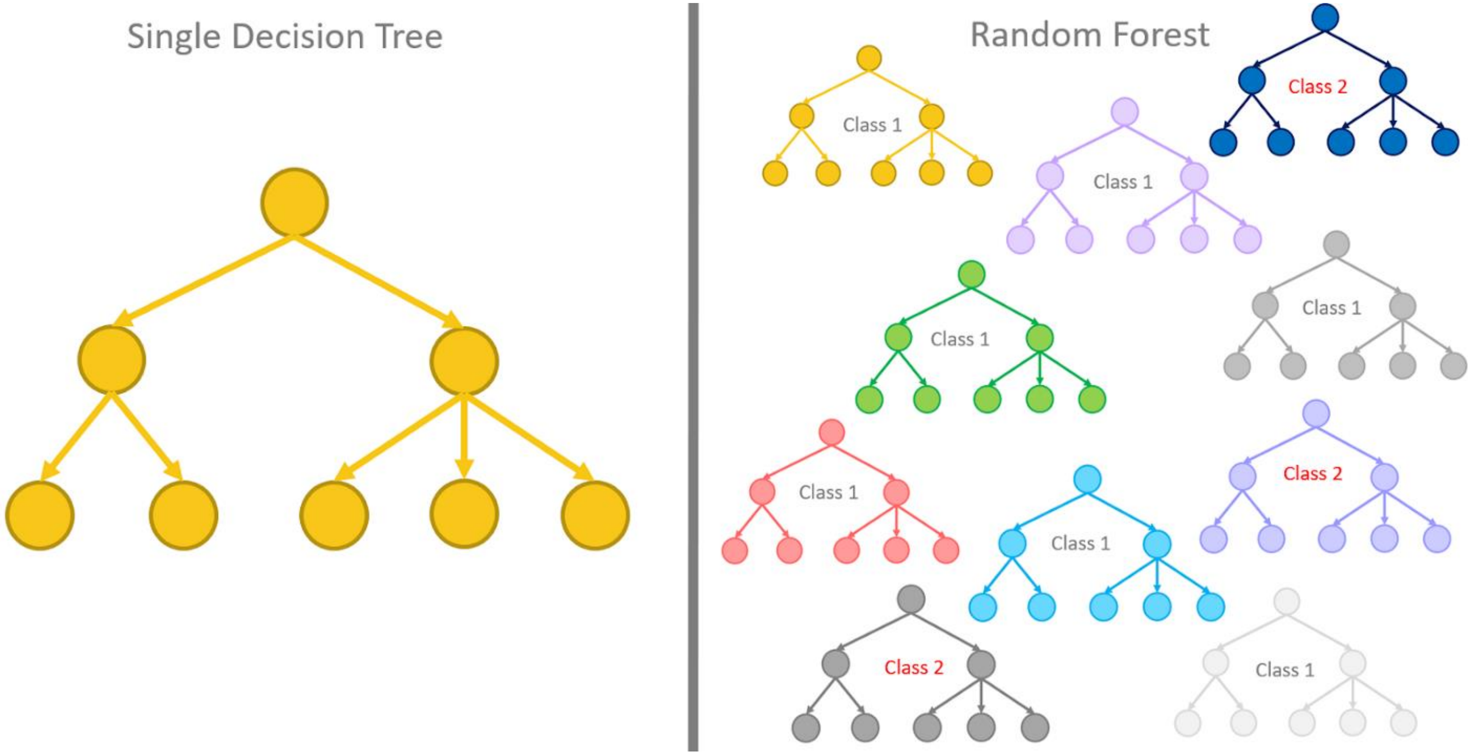


Predictive Analytics Models

Classification:

Classification is the process of creating a set of classes for data based on the existing data.

Decision trees
Random forest



Predictive Analytics Models

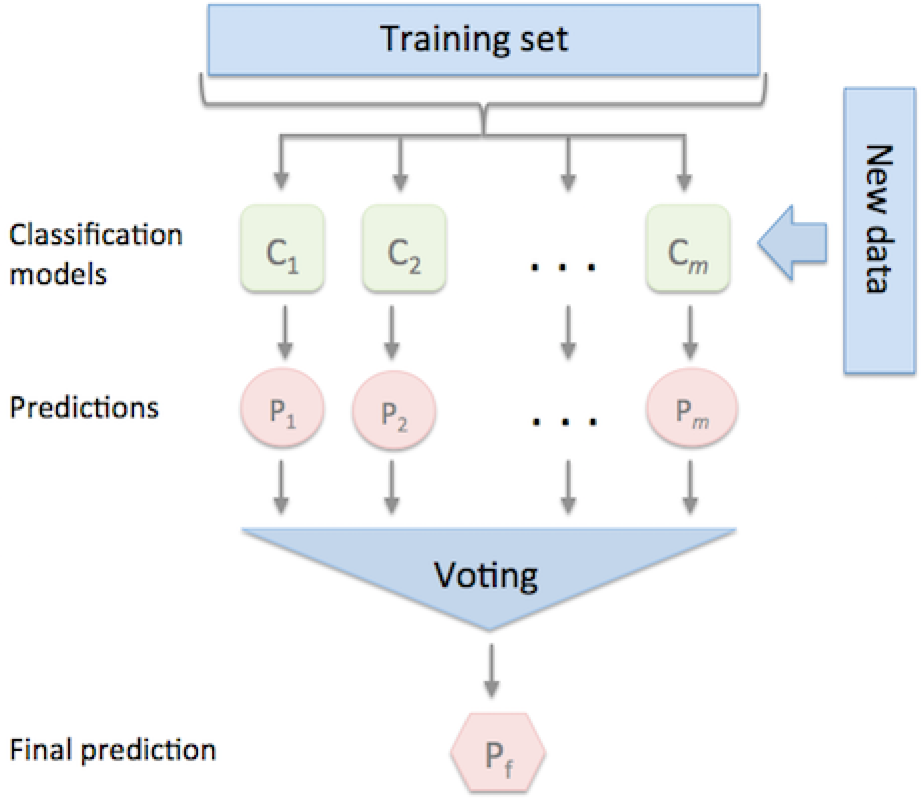
Classification:

Classification is the process of creating a set of classes for data based on the existing data.

Decision trees

Random forest

Voting classifiers

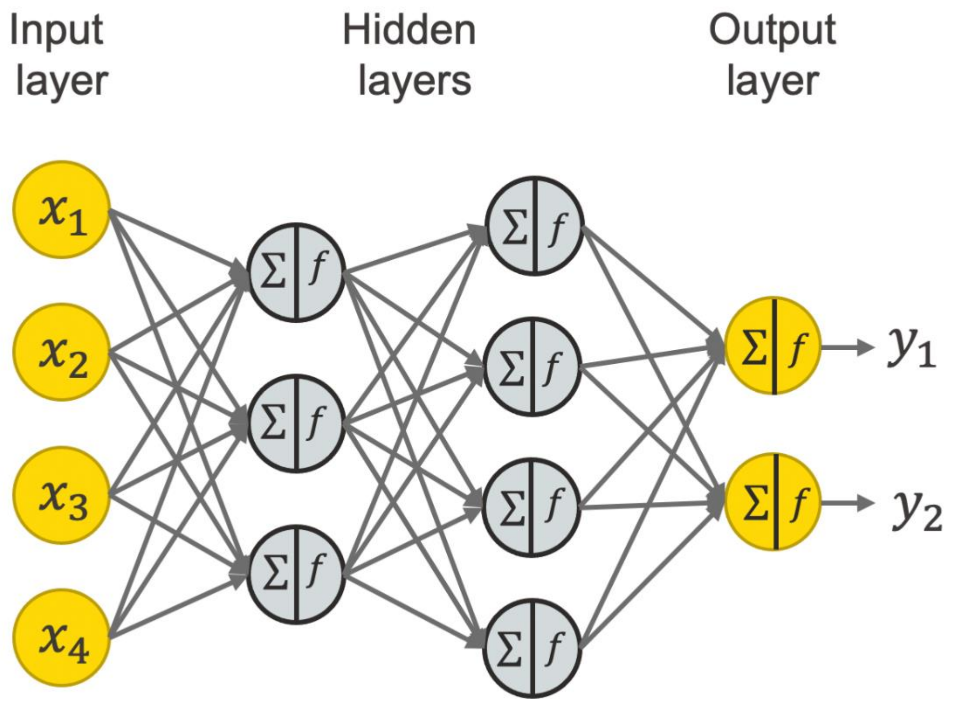


Predictive Analytics Models

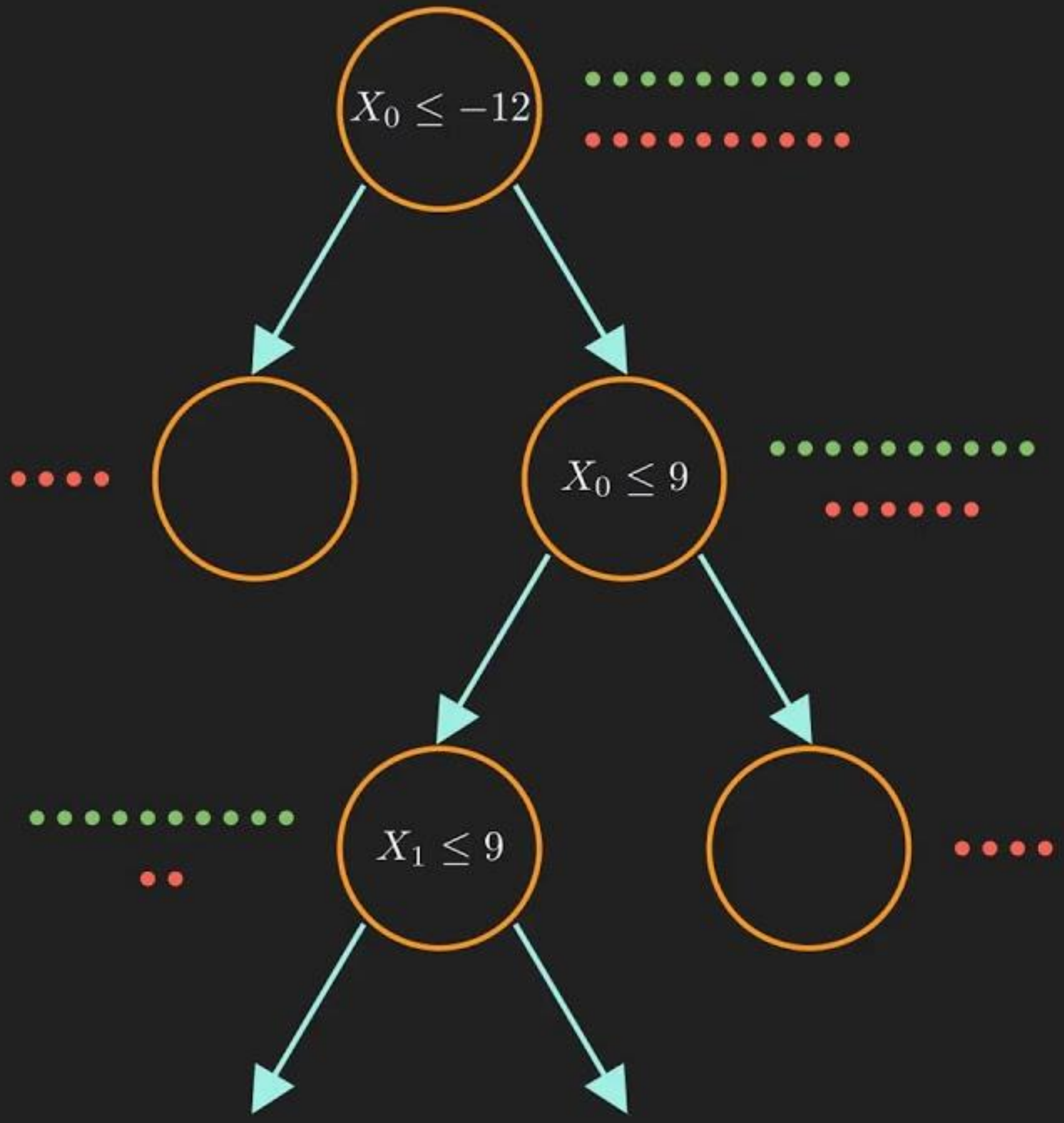
Classification:

Classification is the process of creating a set of classes for data based on the existing data.

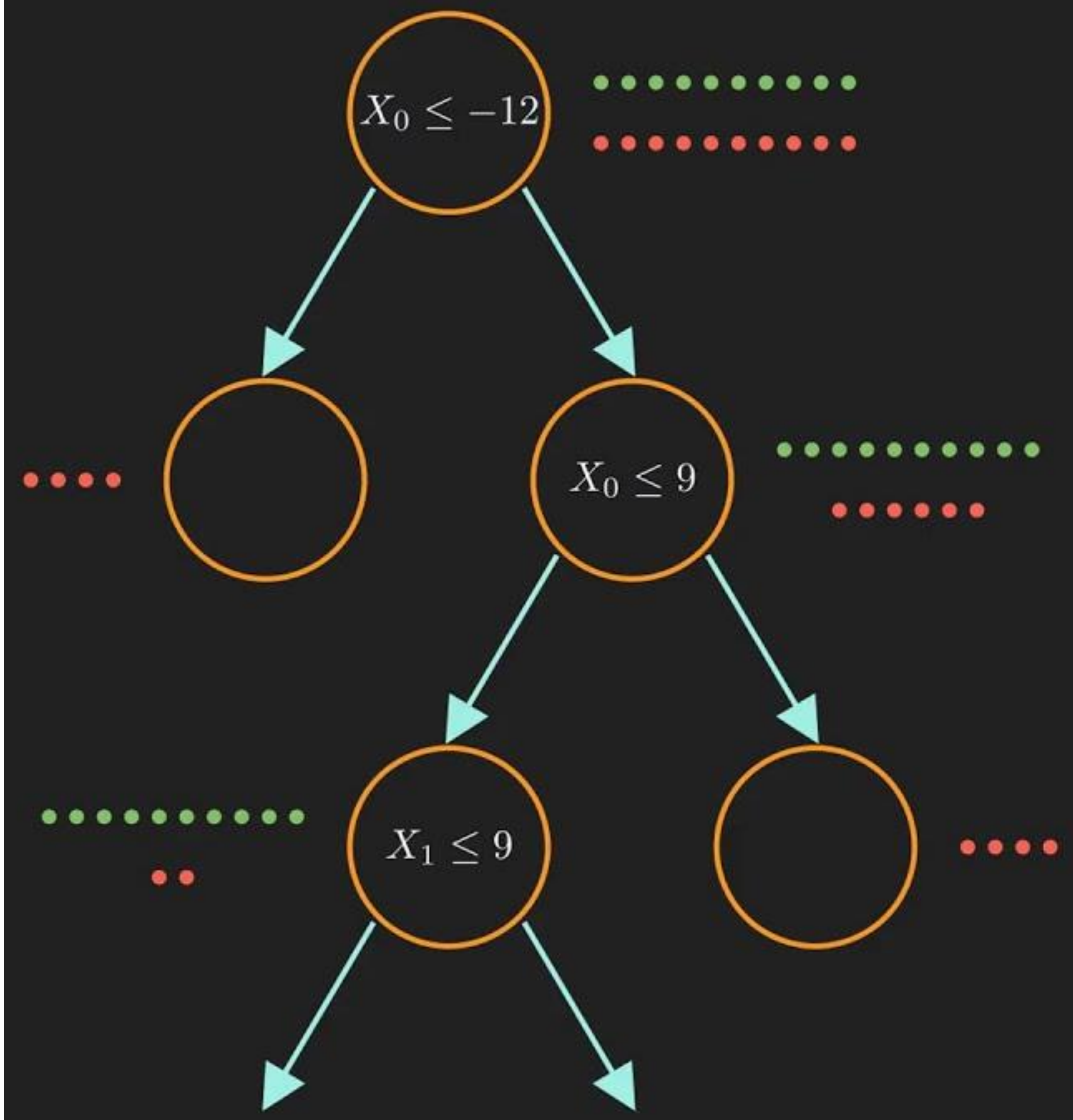
Decision trees
Random forest
Voting classifiers
Neural networks and deep learning
…



Decision Tree Classifier



- A decision tree is essentially **an upside-down tree shaped diagram used to classify.**
- It is a predictive model **based on a branching series of Boolean tests (often) and non-Boolean tests.**
- It has a <u>**root node**</u> which is the **starting point of the decision tree**.
- **Splitting or branching** is the process of dividing a node into **two or more sub-nodes**.
- Nodes have **sub-nodes and leaf/terminal nodes**, which are the ones without a split.
- **Sub-nodes** of a specific node is known as **child nodes** and the node is known as parent node.



features or attributes

decision

Outlook	Temp	Humidity	Windy	Play Tennis?
Sunny	Hot	High	Weak	No
Sunny	Hot	High	Strong	No
Overcast	Hot	High	Weak	Yes
Rain	Mild	High	Weak	Yes
Rain	Cool	Normal	Weak	Yes
Rain	Cool	Normal	Strong	No
Overcast	Cool	Normal	Strong	Yes
Sunny	Mild	High	Weak	No
Sunny	Cool	Normal	Weak	Yes
Rain	Mild	Normal	Weak	Yes
Sunny	Mild	Normal	Strong	Yes
Overcast	Mild	High	Strong	Yes
Overcast	Hot	Normal	Weak	Yes
Rain	Mild	High	Strong	No

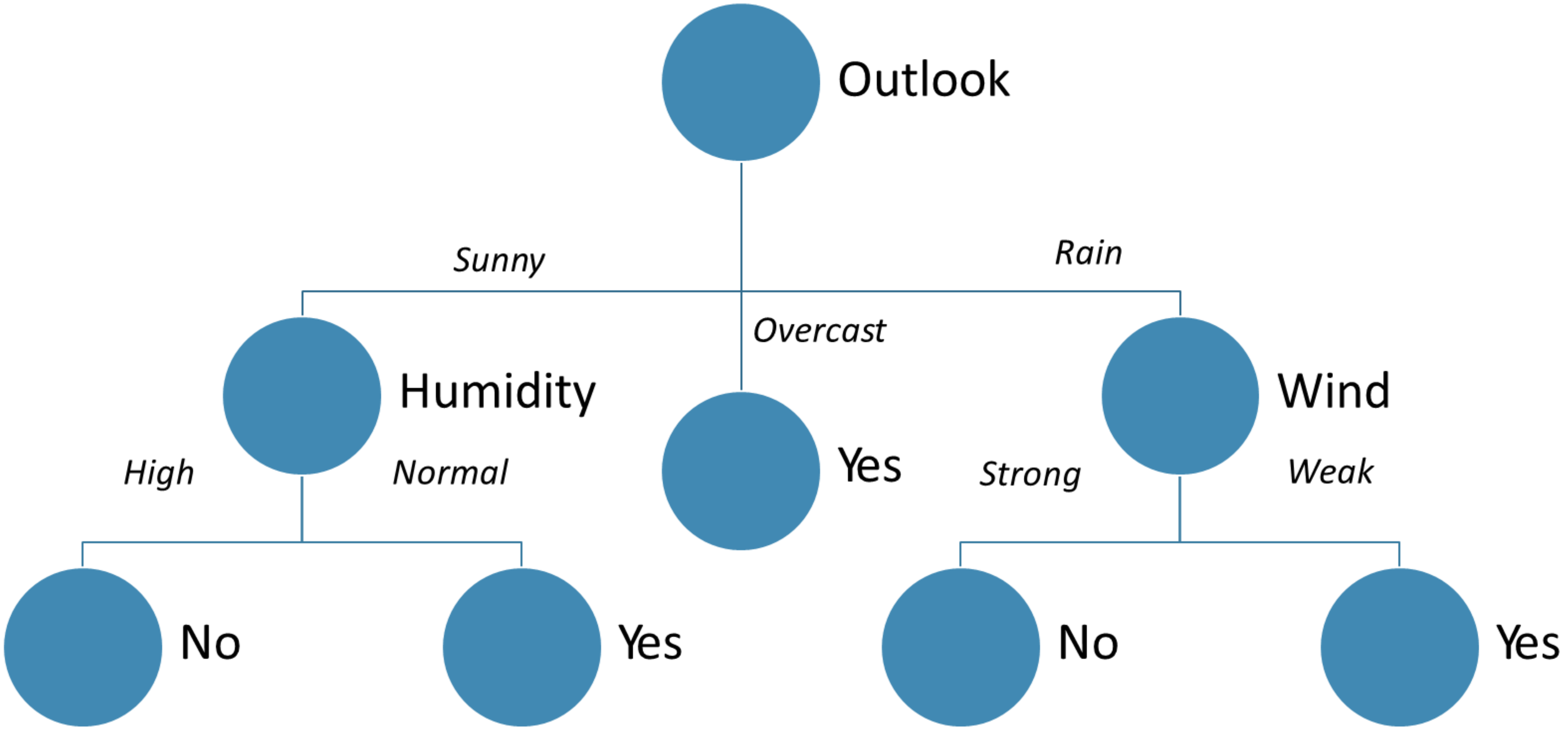
Decision Tree

features or attributes

decision/classes

Outlook	Temp	Humidity	Windy	Play Tennis?
Sunny	Hot	High	Weak	No
Sunny	Hot	High	Strong	No
Overcast	Hot	High	Weak	Yes
Rain	Mild	High	Weak	Yes
Rain	Cool	Normal	Weak	Yes
Rain	Cool	Normal	Strong	No
Overcast	Cool	Normal	Strong	Yes
Sunny	Mild	High	Weak	No
Sunny	Cool	Normal	Weak	Yes
Rain	Mild	Normal	Weak	Yes
Sunny	Mild	Normal	Strong	Yes
Overcast	Mild	High	Strong	Yes
Overcast	Hot	Normal	Weak	Yes
Rain	Mild	High	Strong	No

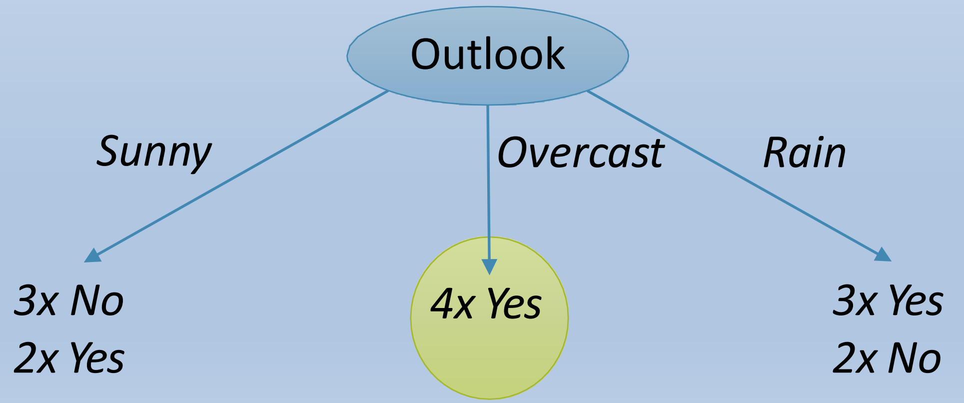
Decision Tree Classification



Decision Tree Classification

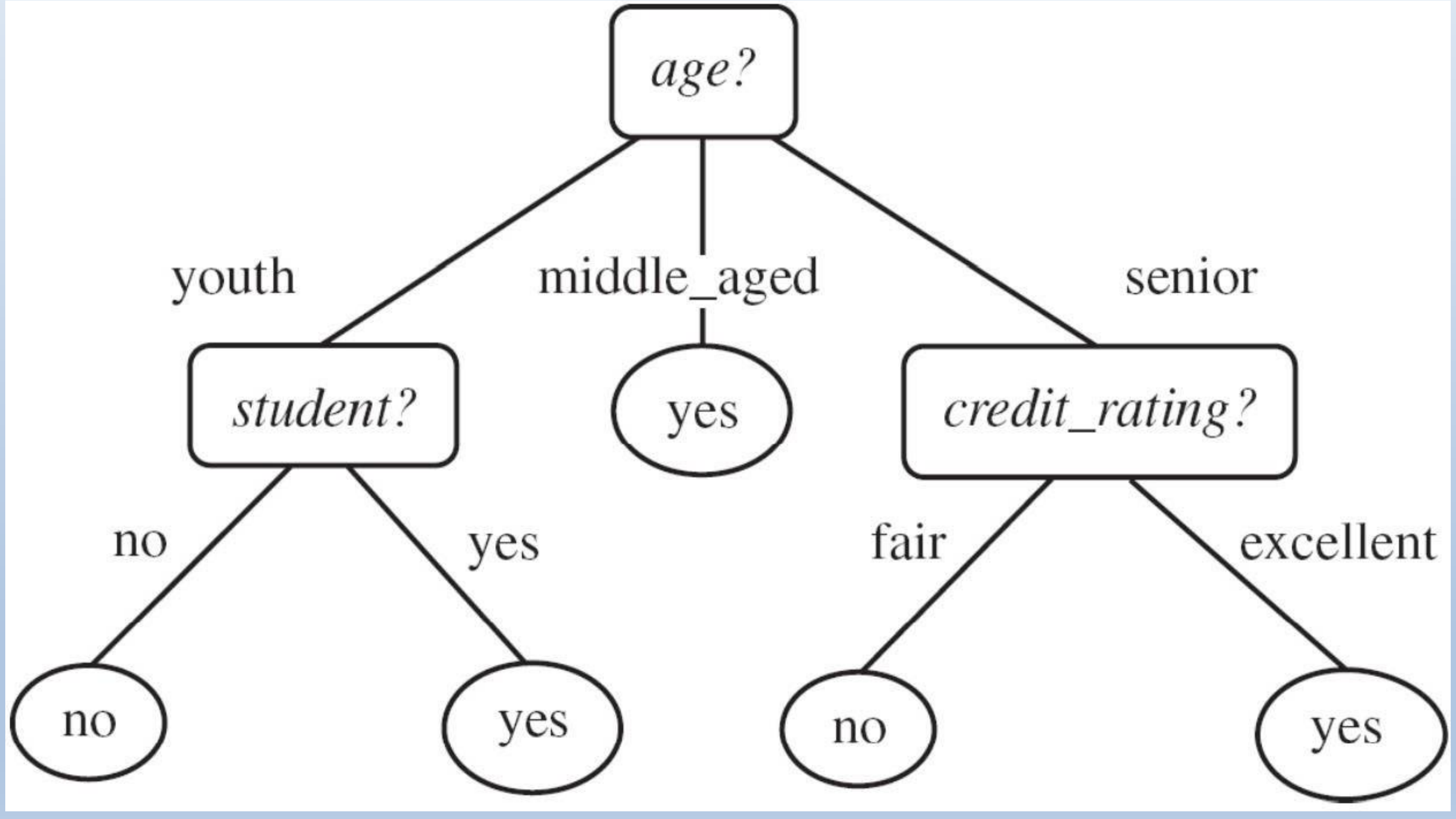
How do you build a decision tree?

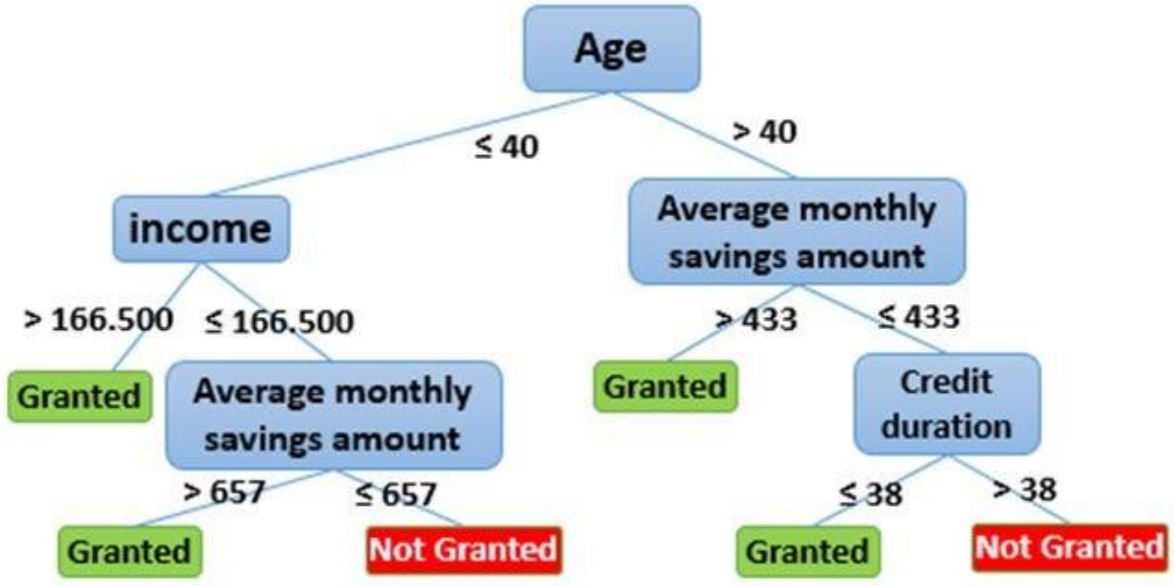
- Select an attribute (A)
  - Which attribute is the best? Impurity or information gain can be used
- For each value of A, create a partition.
- If training samples are perfectly classified, then stop – otherwise recursively iterate over the new child nodes



Example: Classifying potential customers using a decision tree

RID	*age*	*income*	*student*	*credit_rating*	*Class: buys_computer*
1	youth	high	no	fair	no
2	youth	high	no	excellent	no
3	middle_aged	high	no	fair	yes
4	senior	medium	no	fair	yes
5	senior	low	yes	fair	yes
6	senior	low	yes	excellent	no
7	middle_aged	low	yes	excellent	yes
8	youth	medium	no	fair	no
9	youth	low	yes	fair	yes
10	senior	medium	yes	fair	yes
11	youth	medium	yes	excellent	yes
12	middle_aged	medium	no	excellent	yes
13	middle_aged	high	yes	fair	yes
14	senior	medium	no	excellent	no


age?
youth
middle_aged
senior
student?
yes
credit_rating?
no
yes
fair
excellent
no
yes
no
yes



Home Loan Approval

Figure 2
Example Decision Tree:
Should we develop a new
product or consolidate?

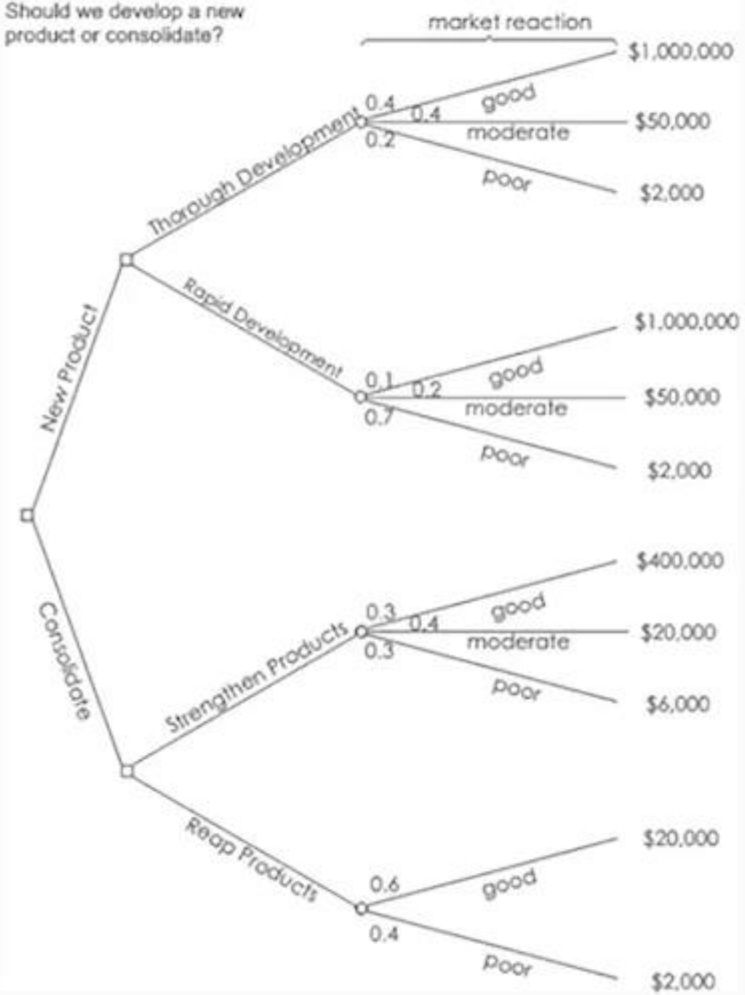
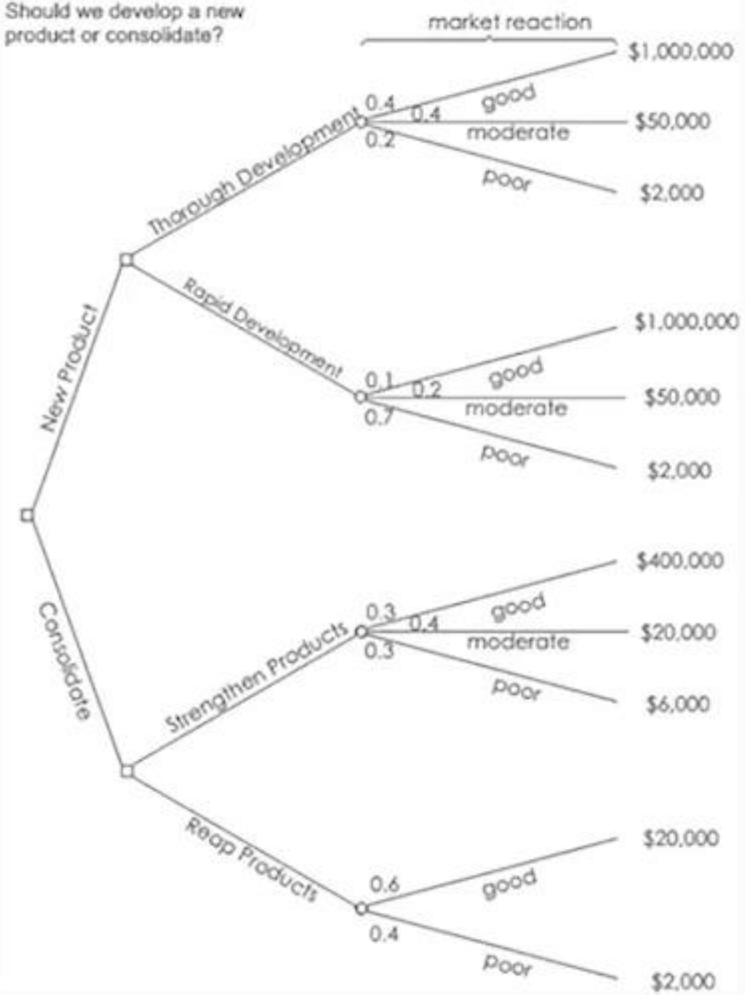


Figure 2
Example Decision Tree:
Should we develop a new
product or consolidate?



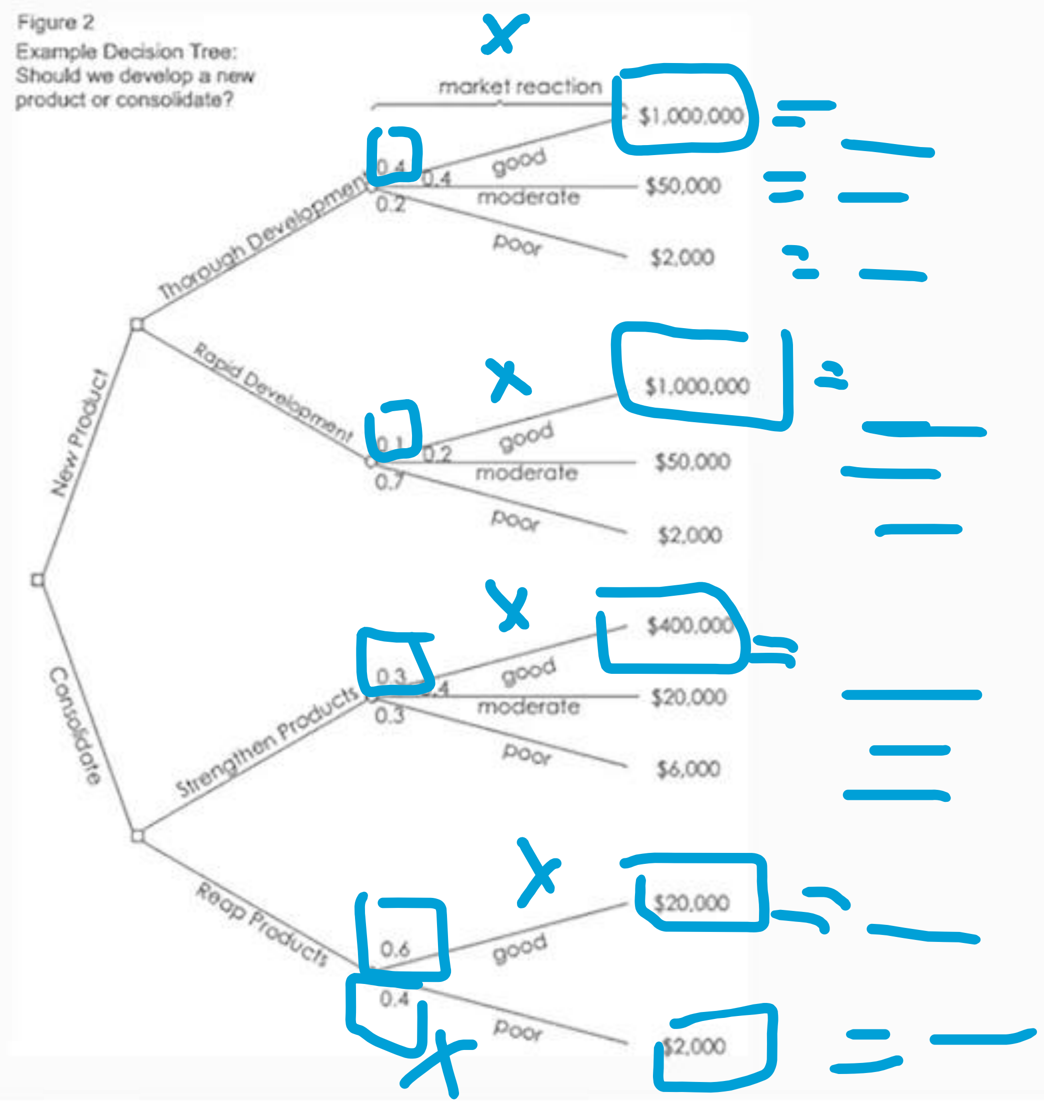
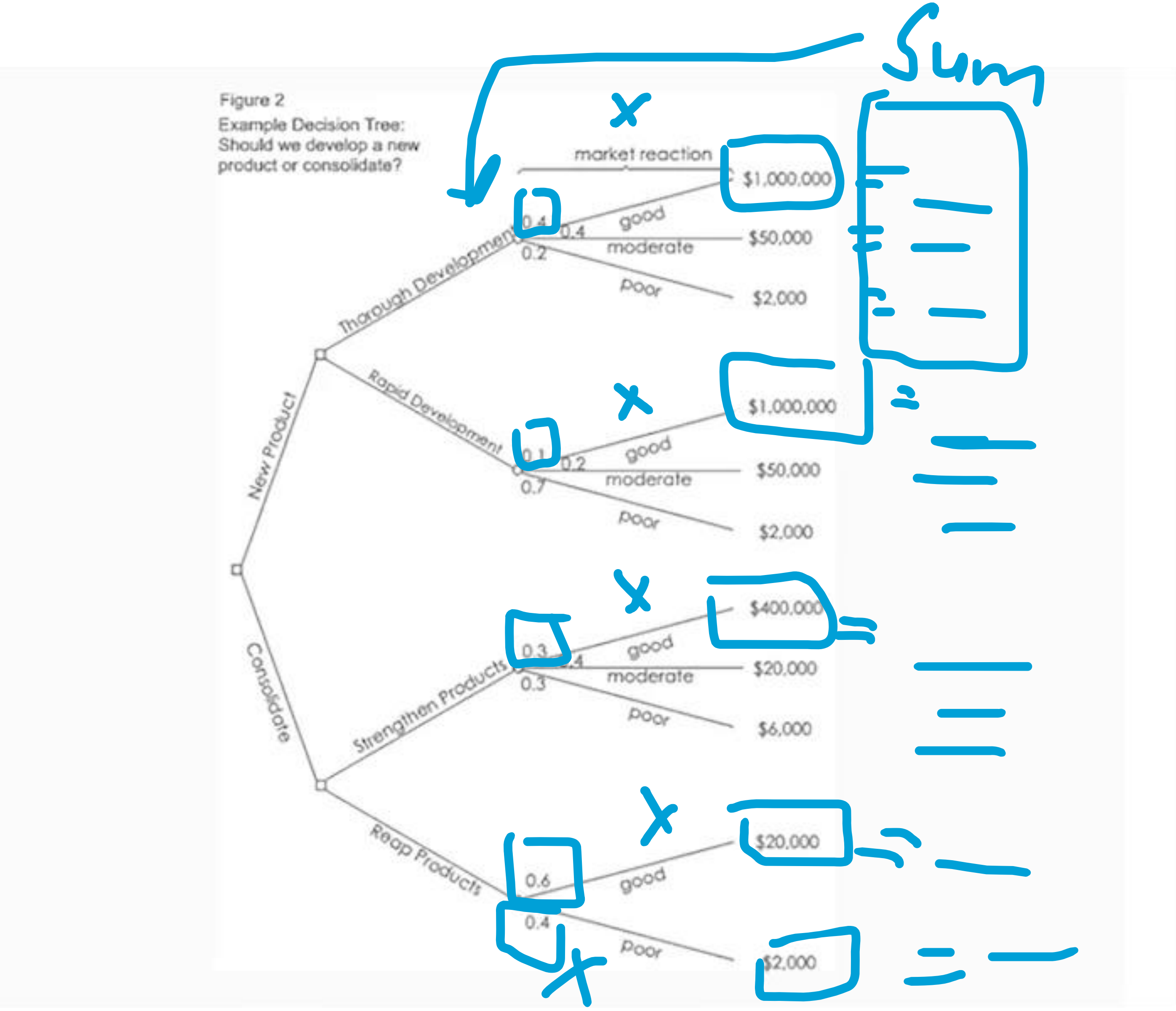


Figure 2
Example Decision Tree:
Should we develop a new product or consolidate?
market reaction
New Product
Thorough Development
Rapid Development
Consolidate
Strengthen Products
Reap Products
good
moderate
poor
0.4
0.4
0.2
0.1
0.2
0.7
0.3
0.4
0.3
0.6
0.4
$1,000,000
$50,000
$2,000
$1,000,000
$50,000
$2,000
$400,000
$20,000
$6,000
$20,000
$2,000



Sum
Figure 2
Example Decision Tree:
Should we develop a new product or consolidate?
market reaction
x
New Product
Thorough Development
0.4
0.4
0.2
good
moderate
poor
$1,000,000
$50,000
$2,000
Rapid Development
0.1
0.2
0.7
good
moderate
poor
$1,000,000
$50,000
$2,000
Consolidate
Strengthen Products
0.3
0.4
0.3
good
moderate
poor
$400,000
$20,000
$6,000
Reap Products
0.6
0.4
good
poor
$20,000
$2,000

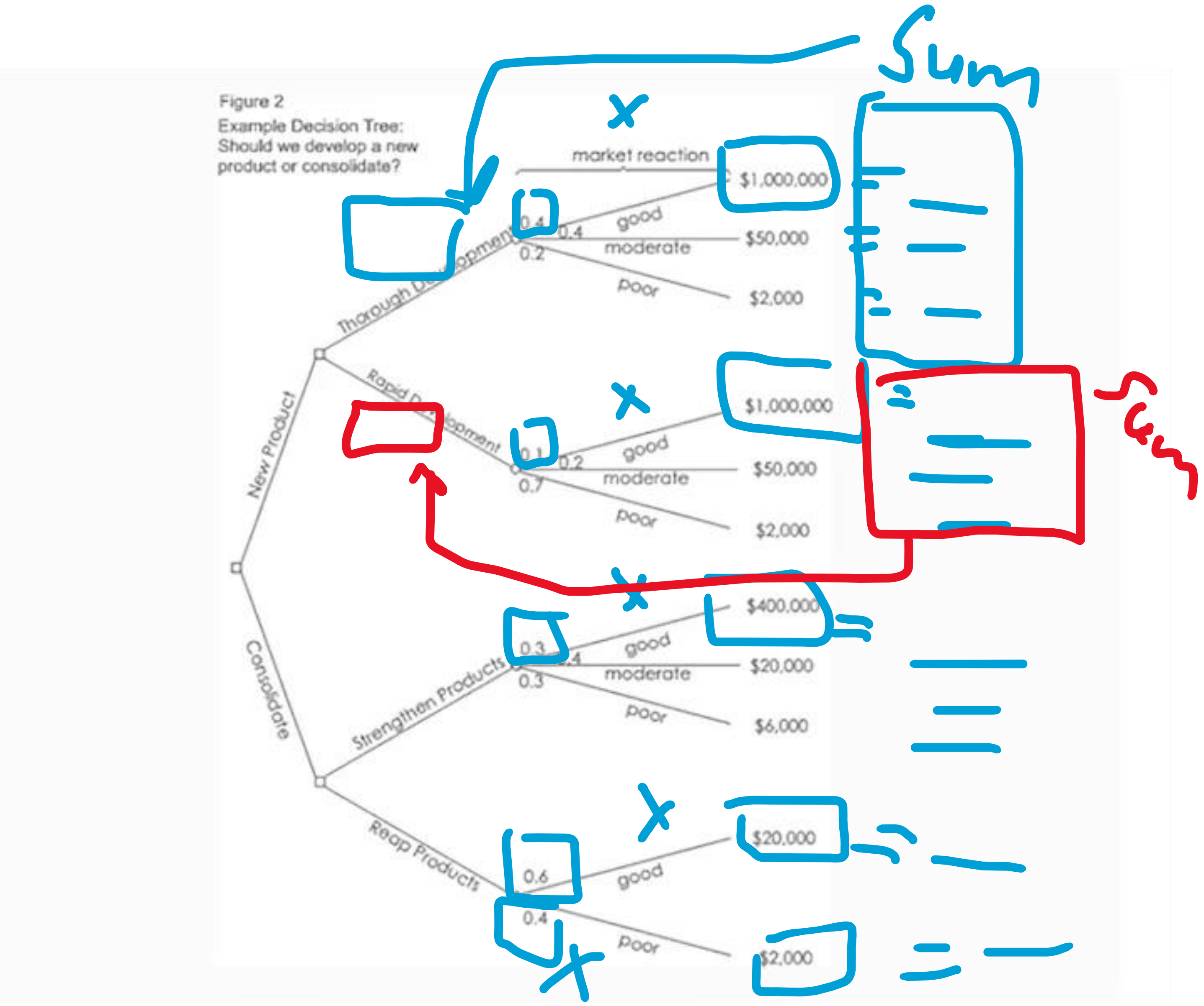


Figure 2
Example Decision Tree:
Should we develop a new product or consolidate?
market reaction
New Product
Thorough Development
Rapid Development
Consolidate
Strengthen Products
Reap Products
0.4
0.4
0.2
good
moderate
poor
$1,000,000
$50,000
$2,000
0.1
0.2
0.7
$1,000,000
$50,000
$2,000
0.3
0.4
0.3
$400,000
$20,000
$6,000
0.6
0.4
good
poor
$20,000
$2,000
Sum
Sum
x
x
x
x
x

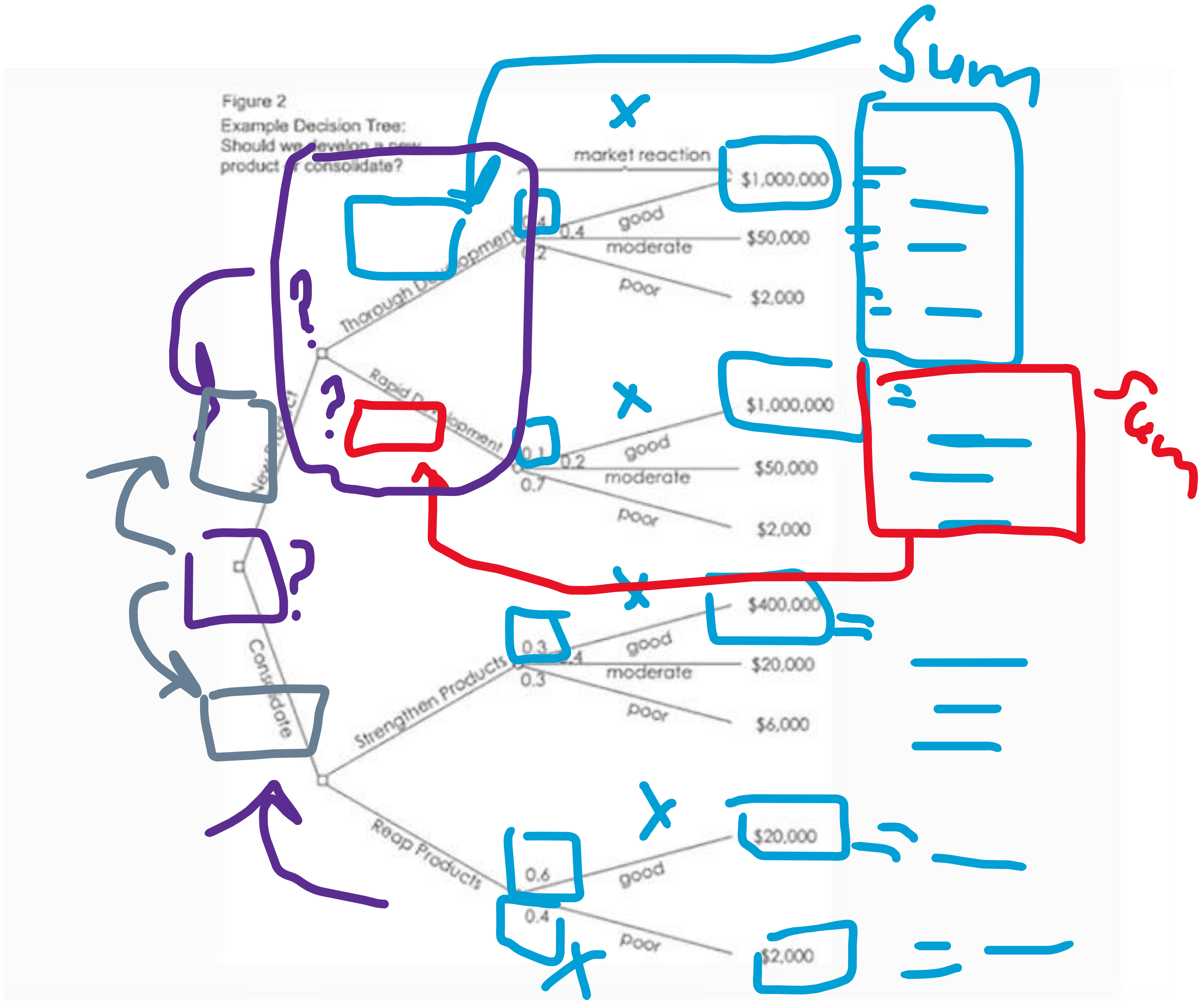


Figure 2
Example Decision Tree:
Should we develop a new product or consolidate?
Sum
x
market reaction
$1,000,000
0.4
good
0.4
moderate
0.2
poor
$50,000
$2,000
Thorough Development
?
Rapid Development
?
0.1
0.2
good
moderate
0.7
poor
$1,000,000
$50,000
$2,000
Sum
New Product
Consolidate
?
$400,000
0.3
0.4
good
moderate
0.3
poor
$20,000
$6,000
Strengthen Products
Reap Products
0.6
good
0.4
poor
$20,000
$2,000

With limited attributes, it is possible to do the graphing of decision tree by hand, in other cases, it may become really time consuming and confusing.

Later in your jobs, if you wanted to use decision tree, remember that there are automated tools to do decision tree for you.

Applications of Classification Techniques

- **Financial markets** – ex: stock price prediction
- **Marketing** – next best offer prediction
- **Web Analytics** – ‘customer like you also purchased x,y,z’
- **Credit Modelling** – will you get approved for a credit card?
- **Medicine** – predicting the likely protein to bind to a virus
- **Social/Political Science** – predicting who wins the next US election
- **Automotive Industry** – Autopilot

Predictive Analytics Models in Supply Chains

- **Classification**

  Classification is a mathematical model that can differentiate **between two or more outcomes**. For instance, using the historical data of a supplier, and setting some decision attributes, the decision of ‘whether to sign a contract with this supplier?’ can be answered by a classification model. Managers can be provided with such classification models to assist their decisions.

- **Regression**

  As opposed to predicting a decision, regression focus on **predicting an unknown future value using available data**. For instance, given the bitcoin prices for the last year, ‘what will be the price next week?’ can be answered by a regression model. Managers can be provided with the predicted value of these models to assist their decisions.

- **Other models (Clustering, Time Series, Forecasting, and similar)**

Chapter 1: Classification Models

Chapter 2: Learning how to Formulate a Regression Model

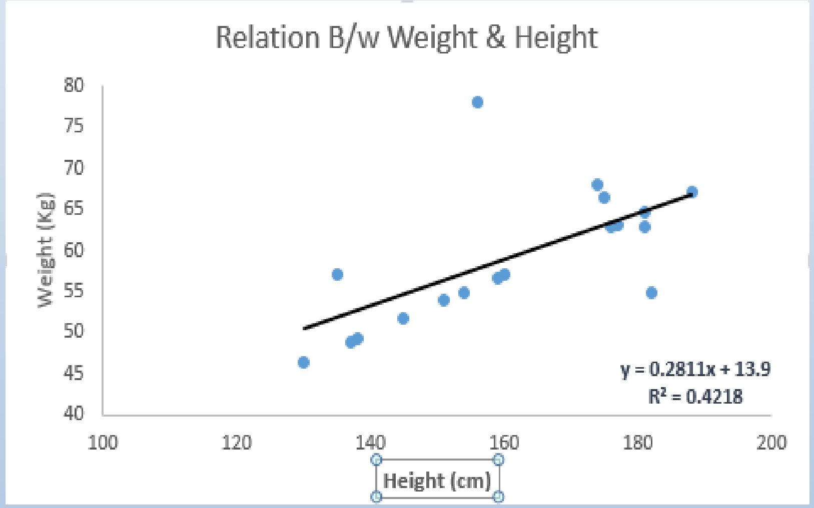
Chapter 3: Utilising Regression Models to Forecast Demand

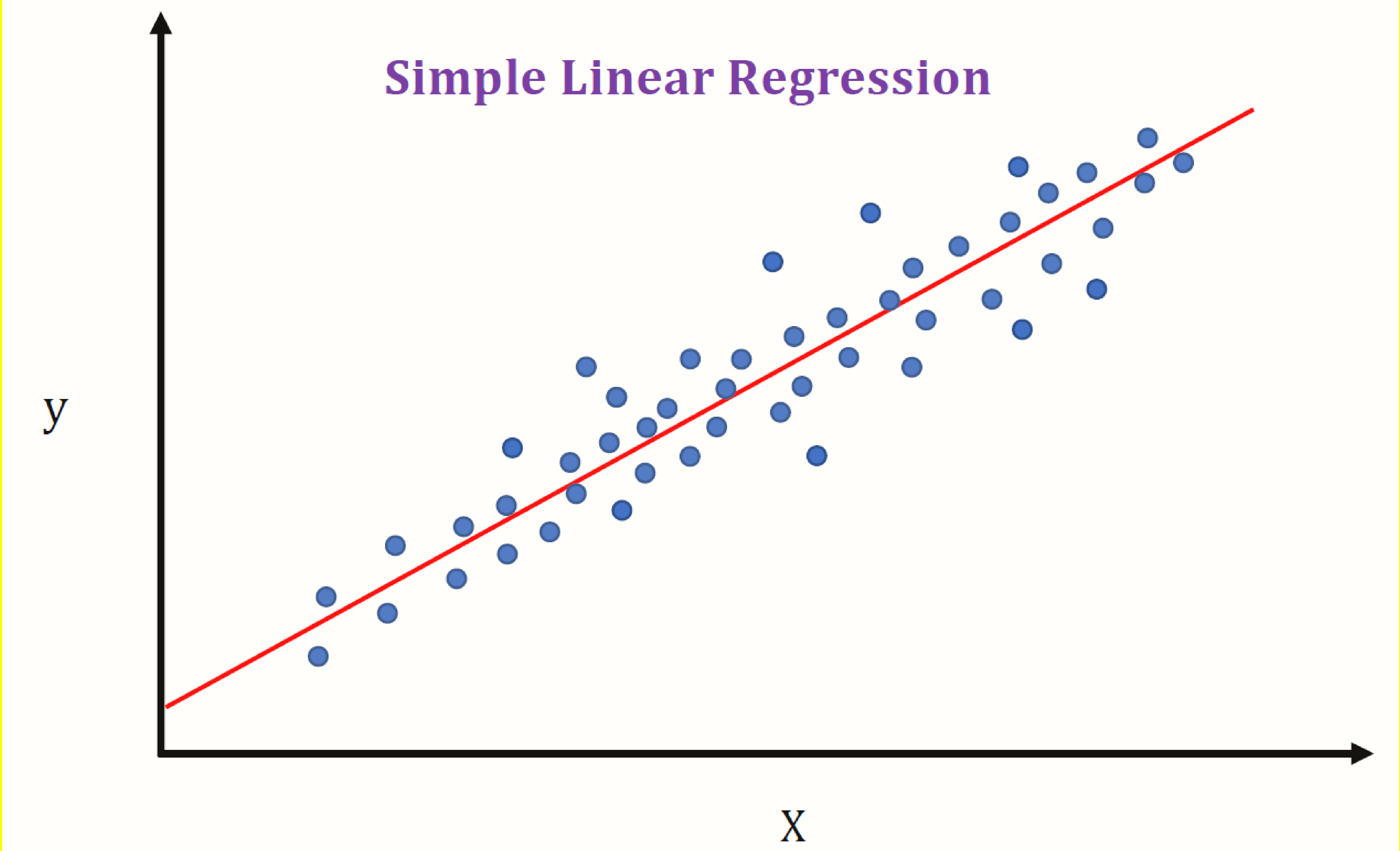
Regression

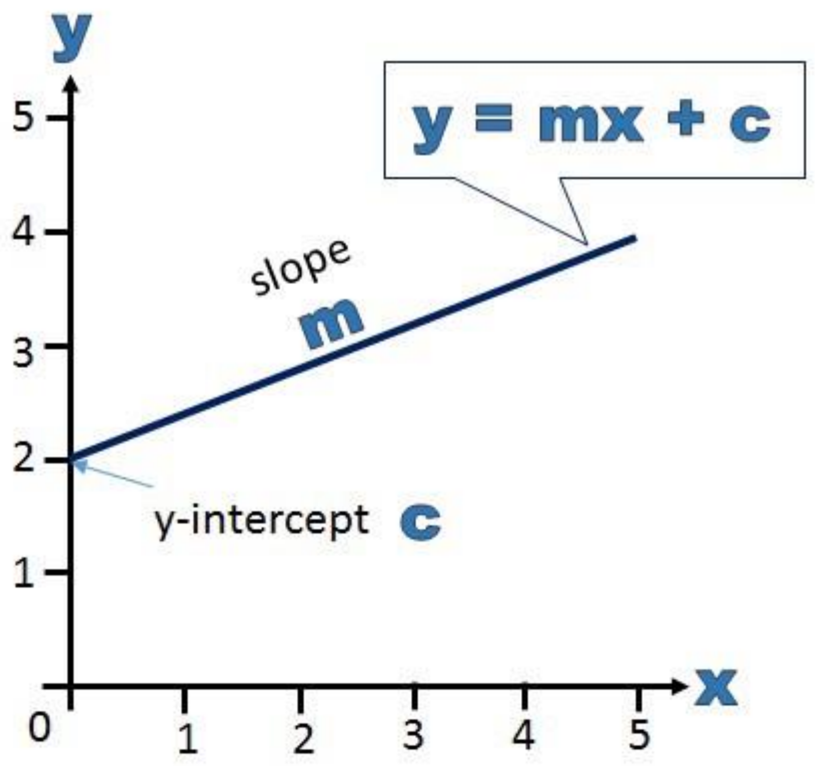
- Regression investigates the relationship between **a dependent (target)** variable and **independent variables (predictor)**
- It is used for forecasting, time series analysis and finding causal effect relationship between variables

Linear Regression

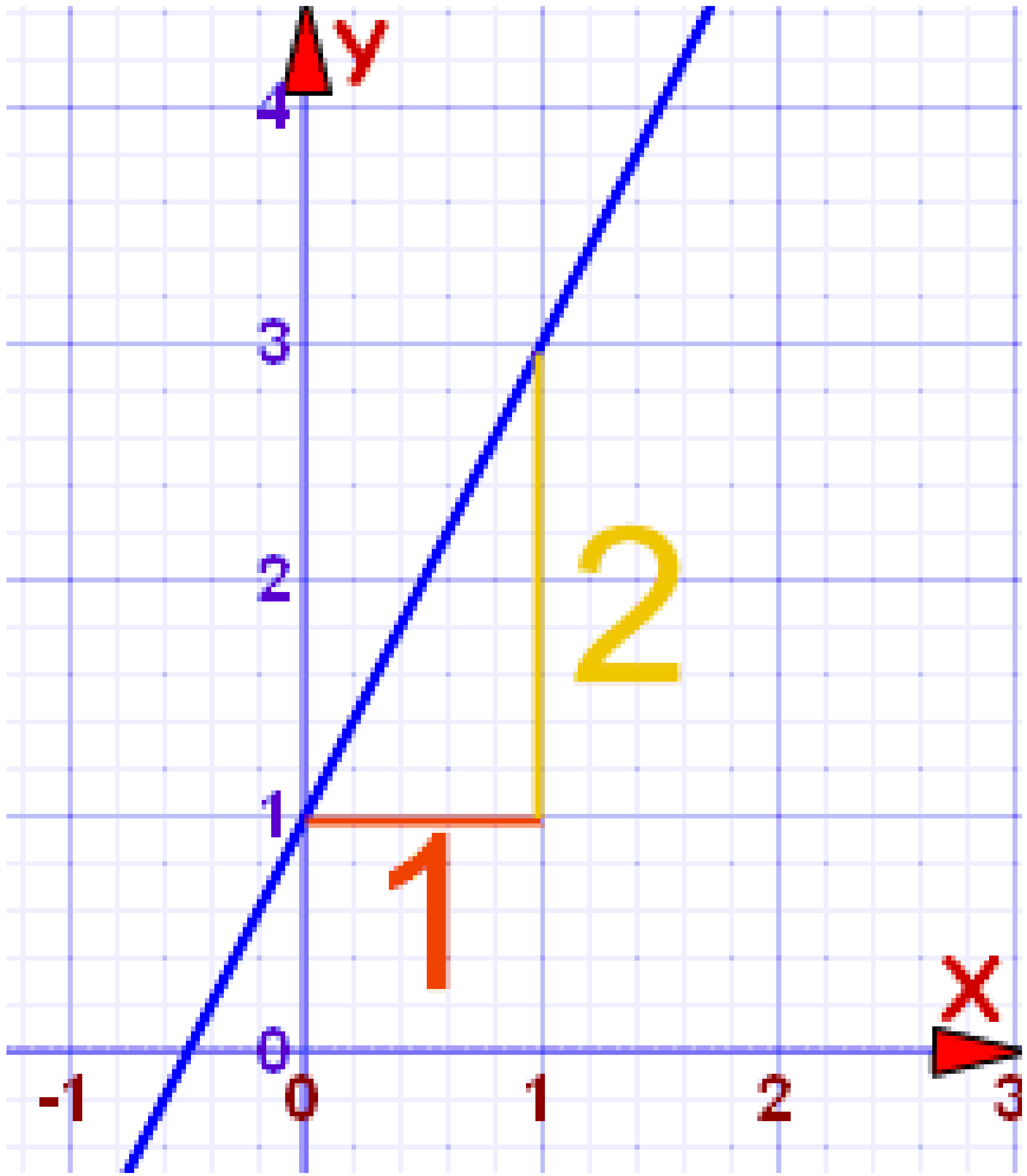
- Establishes a **relationship between dependent variable (y) and one or more independent variables (X)** using best fit straight line (also known as regression line)
- This is represented by an **equation y = a*x + b**, where *a* is the intercept, *b* is the slope of the line and *e* is the error term
- The **available data is graphed to find this regression line. This enables determining future values.** The equation can then be found to predict the value of the target variable based on the independent variables




Simple Linear Regression
y
X


y
y = mx + c
slope
m
y-intercept c
x
0
1
2
3
4
5
1
2
3
4
5

Y=2x+1



Example 1: The relationship between experience (years with company) and salary

- **Click** — Click Data->Data Analysis->Regression
- **Select** — Select your dependent variable for Y range and independent variable for X range.
- **Look** — Look for the value of R2 as it tells us how much change in Y is driven by X.
- **Look** — Look for p-value and coefficients as well

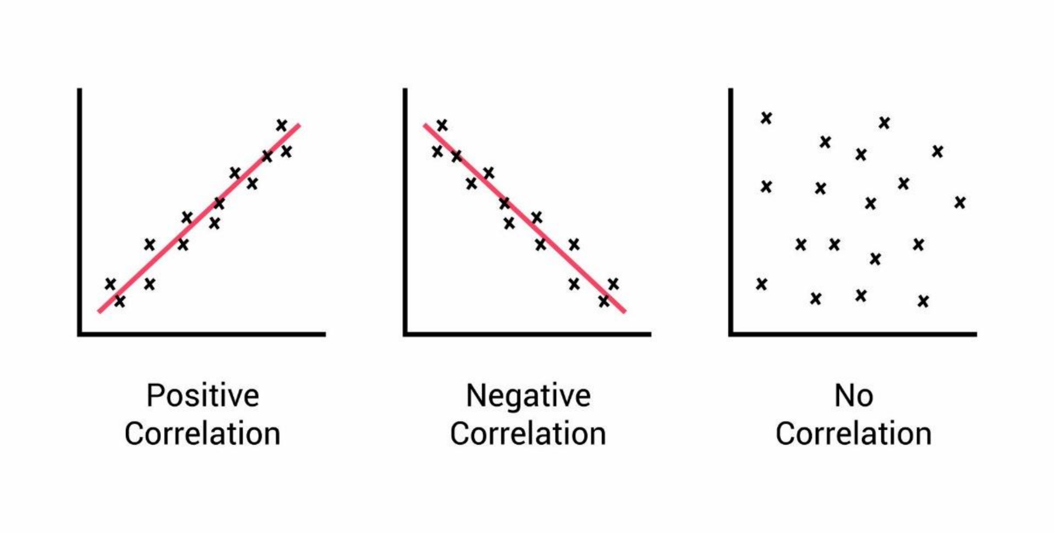
SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.766051647
R Square	0.586835127
Adjusted R Square	0.583883949
Standard Error	15568.50444
Observations	142

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	48196392870	48196392870	198.8477796	1.17585E-28
Residual	140	33932966285	242378330.6		
Total	141	82129359155			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	69816.58422	3023.145185	23.09402293	1.35128E-49	63839.66379	75793.50466	63839.66379	75793.50466
Year with Company	12185.36839	864.1284236	14.10133964	1.17585E-28	10476.94008	13893.79671	10476.94008	13893.79671


Positive
Correlation
Negative
Correlation
No
Correlation

R-squared values determines the proportion of variance in the dependent variable that can be explained by the independent variable. It range from 0 to 1 and are commonly stated as percentages from 0% to 100%.

R-squared values determines the proportion of variance in the dependent variable that can be explained by the independent variable. It range from 0 to 1 and are commonly stated as percentages from 0% to 100%.

A R-squared between 0.50 to 0.99 is acceptable

Similar to p-Value in diagnostic analytics, here, in regression, Significance F needs to be below 0.05 to proceed

Multiple Regression

- This is determining the relationship between **multiple independent variables** and **a dependent variable**. The dependent variable is modelled as a function of several independent variables with corresponding coefficients, along with the constant term.
- Multiple regression requires **two or more independent variables** which is why it's called multiple regression.
- It can be represented by:

$$Y = a_1x_1 + a_2x_2 + \ldots + a_nx_n + b$$

Write the price sale formula

Y: The dependent variable is the price sale
X1, X2: The independent variables are manufacturing and inventory costs

The rest are fixed costs (your intercept value)

F	G
Regression Statistics	
Multiple R	0.880587144
R Square	0.775433719
Adjusted R Square	0.725530101
Standard Error	9111.869683
Observations	12
ANOVA	
	df
Regression	2
Residual	9
Total	11
	Coefficients
Intercept	16387.30
Production Cost	2.03
Inventory Cost	3.58

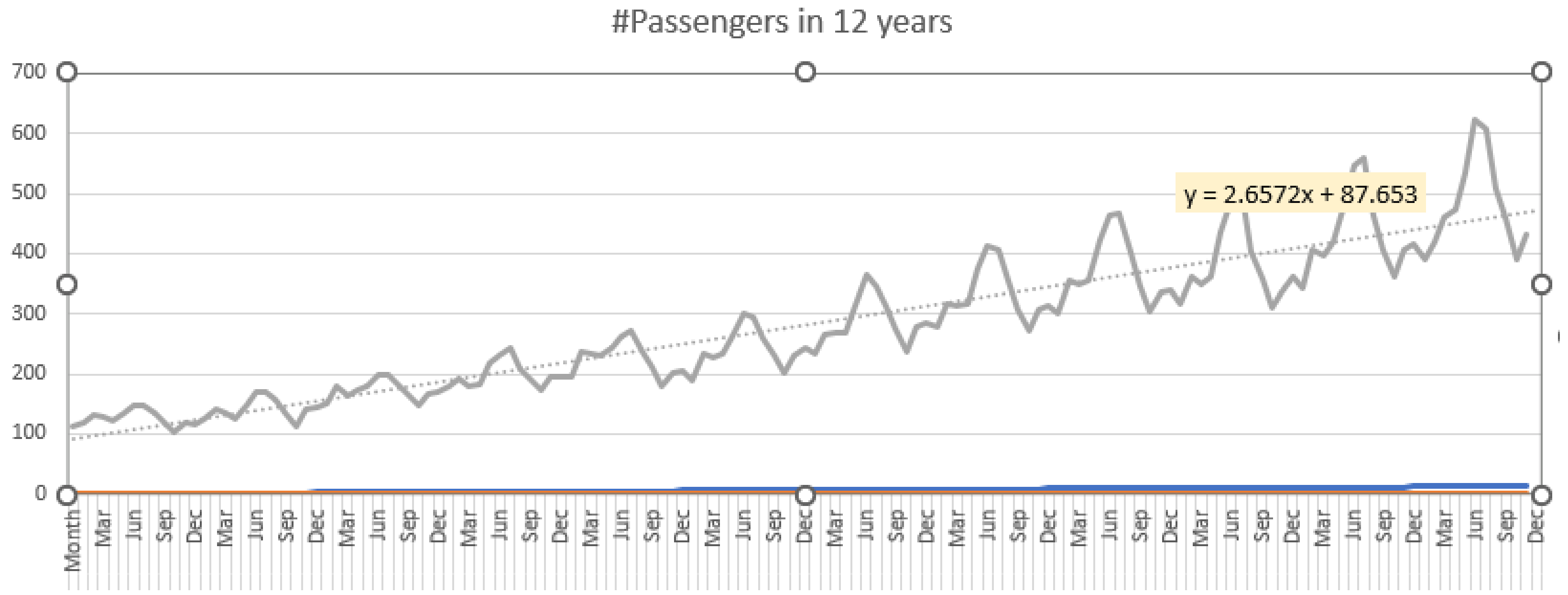
Product price= a * Production Cost + b * Inventory cost + c

Chapter 1: Classification Models

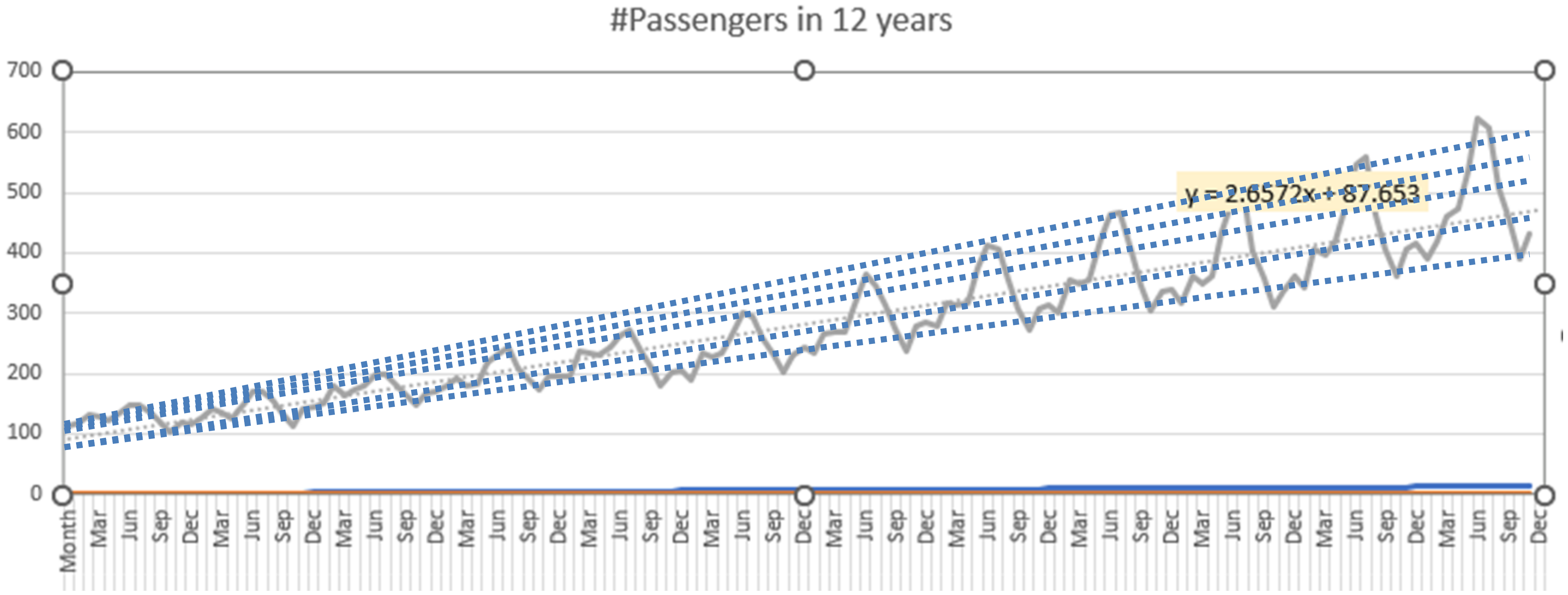
Chapter 2: Learning how to Formulate a Regression Model

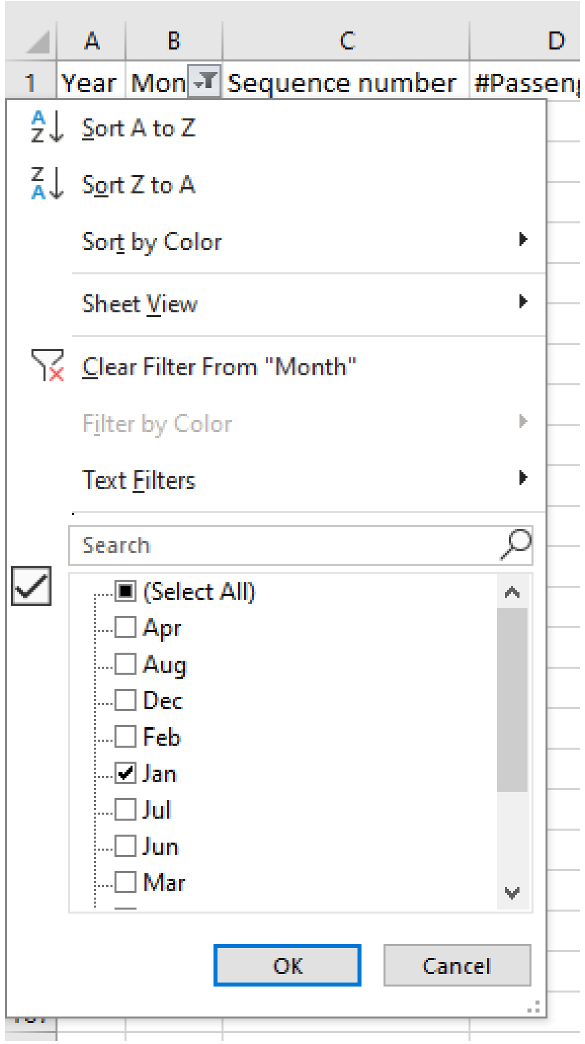
Chapter 3: Utilising Regression Models to Forecast Demand

These values do not consider the seasonality (historical data says that it is always above the trend in July and below the trend in Oct). These predictions are on the trend



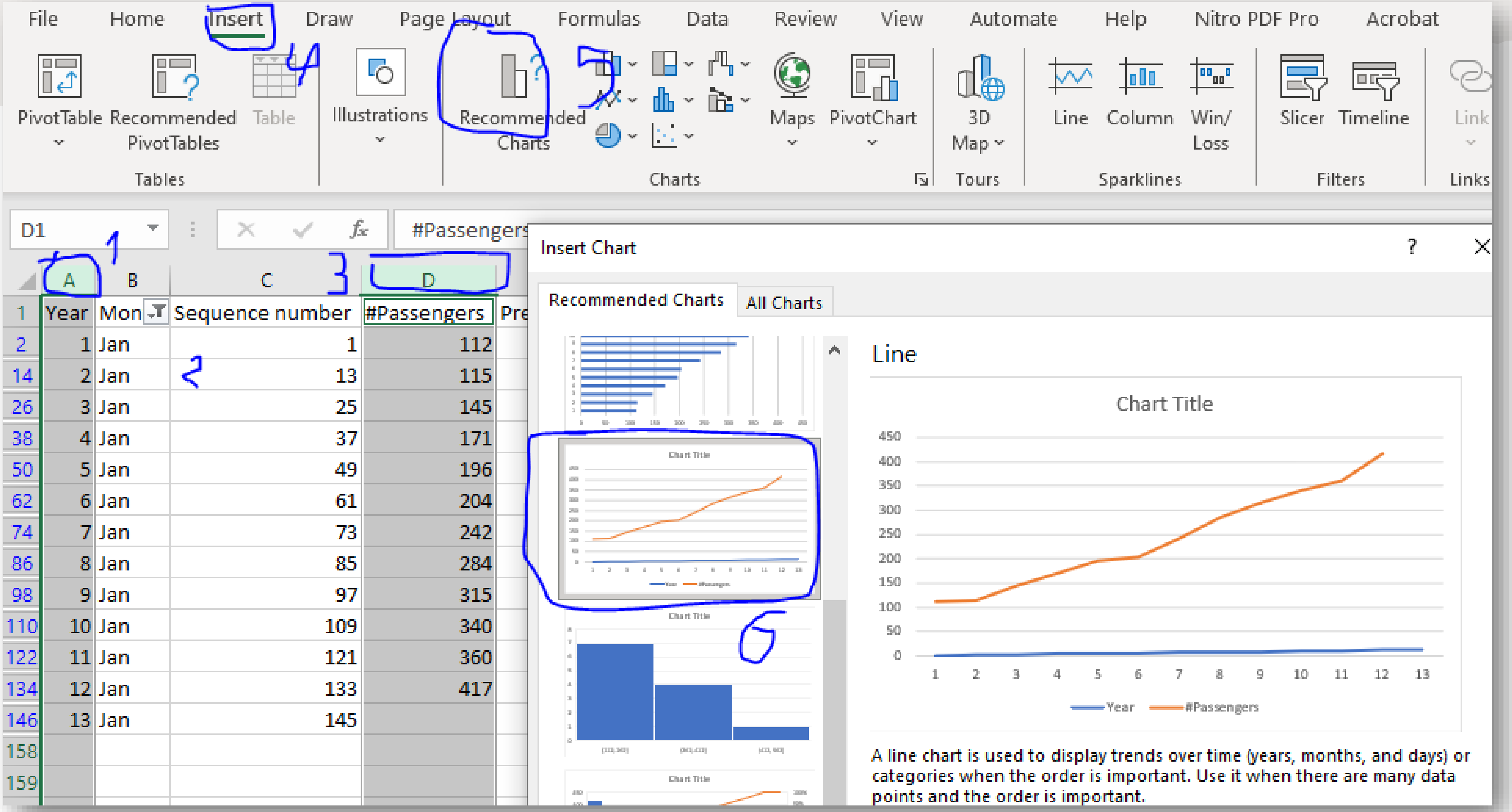
It is always more than the linear trend in specific months of each year, and below the linear trend in some other specific months





A
B
C
D
1
Year
Mon
Sequence number
#Passeng
Sort A to Z
Sort Z to A
Sort by Color
Sheet View
Clear Filter From "Month"
Filter by Color
Text Filters
Search
(Select All)
Apr
Aug
Dec
Feb
Jan
Jul
Jun
Mar
OK
Cancel

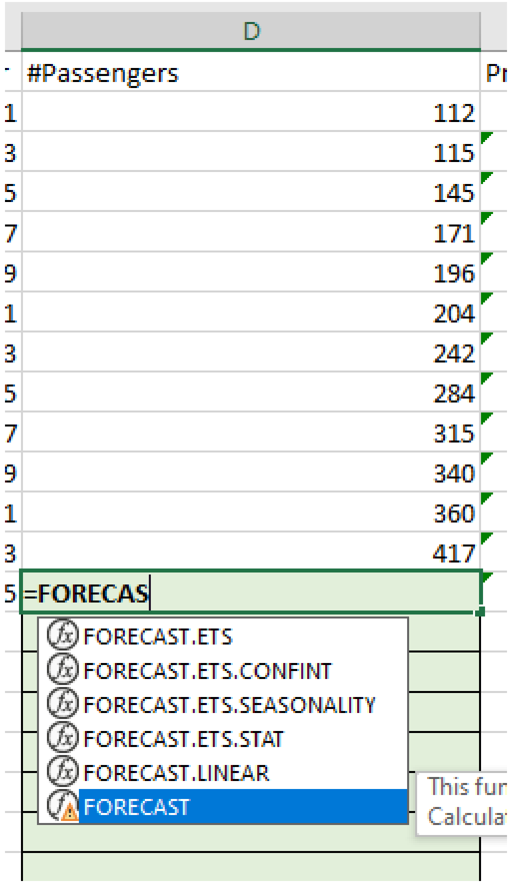
Select YEAR and #Passengers and draw a line chart



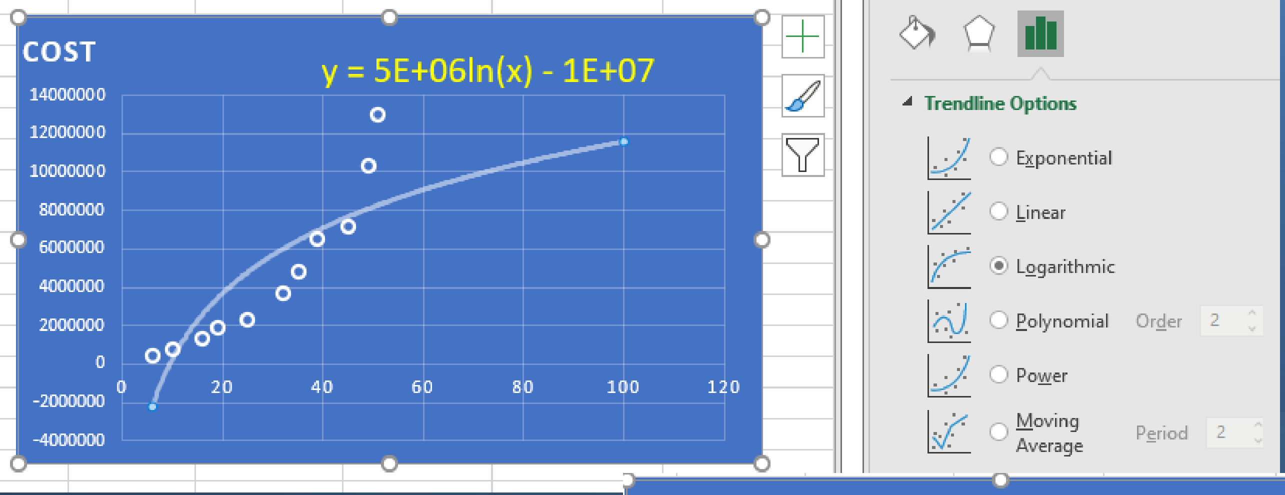
What is next?

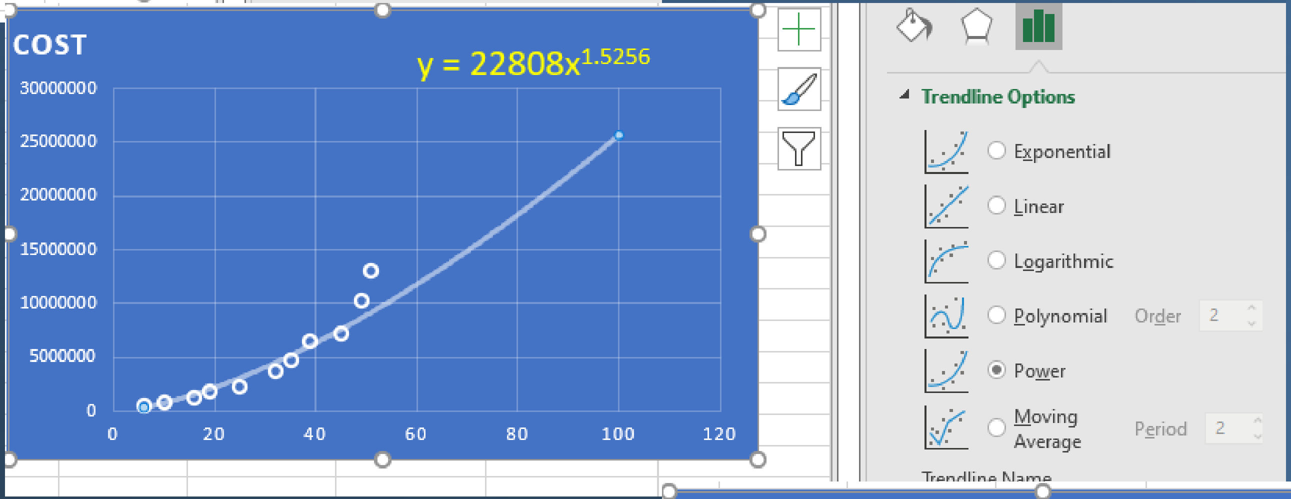
You repeat the last few steps that you did (to find predicted values for Jan year 13) for 11 more times, to compute values for Feb, March, April…. Dec of year 13

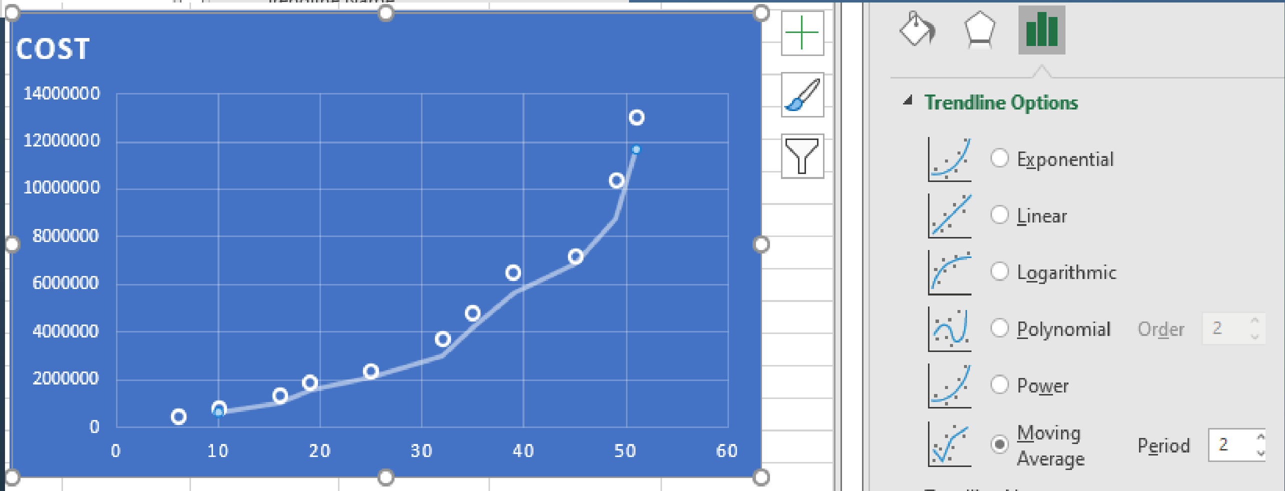
Using Forecast Function – to predict value for Jan

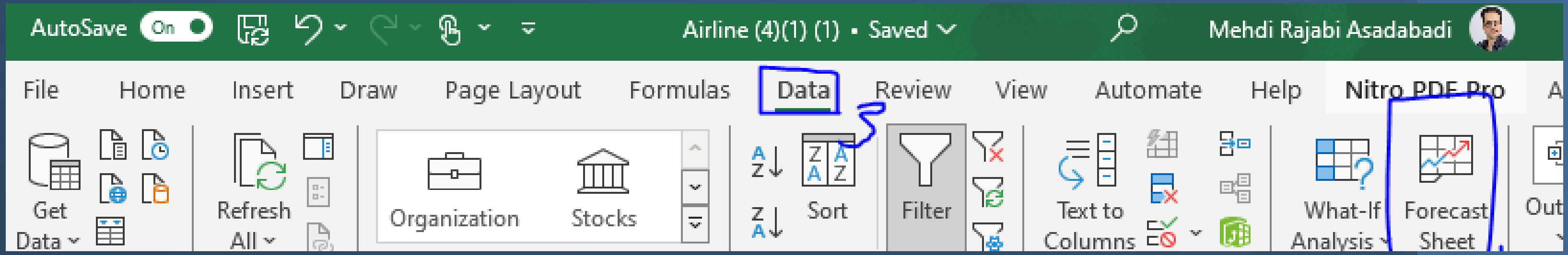


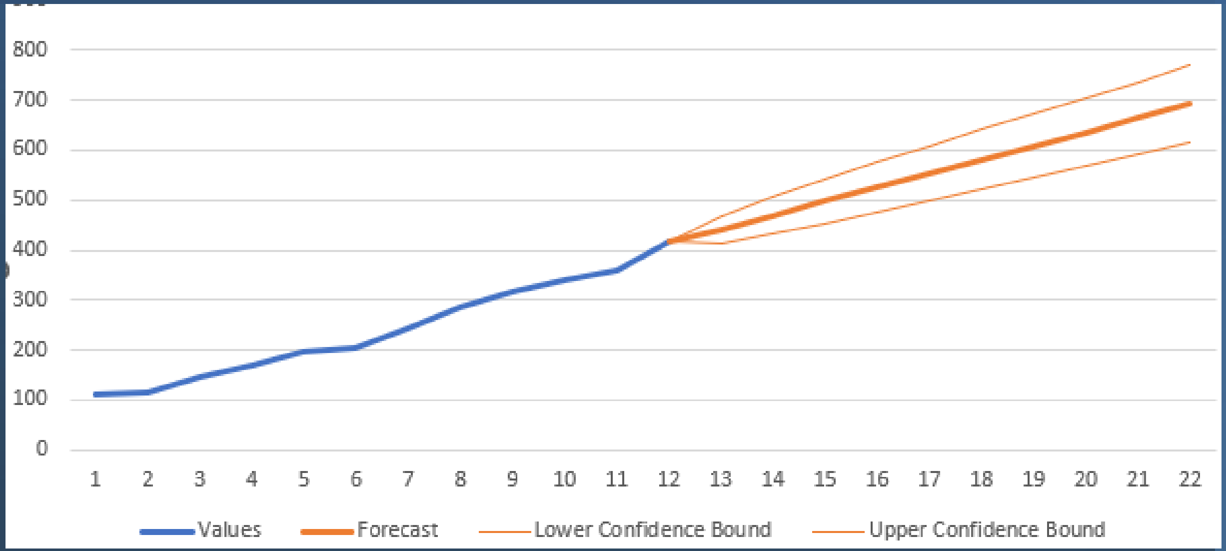
What about non-linear models???


COST
y = 5E+06ln(x) - 1E+07
Trendline Options
Exponential
Linear
Logarithmic
Polynomial
Order
Power
Moving Average
Period


COST
y = 22808x^1.5256
Trendline Options
Exponential
Linear
Logarithmic
Polynomial
Order
Power
Moving Average
Period


COST
Trendline Options
Exponential
Linear
Logarithmic
Polynomial
Order
Power
Moving Average
Period


AutoSave
On
Airline (4)(1) (1) • Saved
Mehdi Rajabi Asadabadi
File
Home
Insert
Draw
Page Layout
Formulas
Data
Review
View
Automate
Help
Nitro PDF Pro
Get Data
Refresh All
Organization
Stocks
Sort
Filter
Text to Columns
What-If Analysis
Forecast Sheet



800
700
600
500
400
300
200
100
0
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22
Values
Forecast
Lower Confidence Bound
Upper Confidence Bound

1. Predict the inventory levels needed for the next quarter.
   Use historical inventory data trends and sales forecasts to estimate future inventory requirements.
2. Forecast inventory costs for the next month.
   Based on historical 'Inventory Cost Per Unit' data, predict the upcoming month's inventory costs.
3. Estimate future stock replenishment needs for high-demand products.
   Use sales trends to predict when high-demand products will need restocking.
4. Predict which products are at risk of stockouts.
   Analyse past inventory levels and sales data to identify products that might face stockouts.
5. Model the impact of a proposed discount strategy on inventory levels.
   Based on past sales and inventory data, simulate how new discounts might affect future inventory needs.

Examples of Predictive Analytics in Inventory Data

1. Forecast the sales trend for the next quarter.
   Use time series forecasting on 'Units Sold' to predict future sales trends.
2. Predict the effect of a 10% discount on high-selling products.
   Model potential impact of a 10% discount on high-selling products based on past data.
3. Estimate future sales during peak seasonal periods.
   Analyze past seasonal sales to predict future sales during peak seasons.
4. Predict changes in supplier quality ratings over time.
   Model potential trends in quality ratings based on industry dynamics.
5. Forecast the potential impact of supplier cost changes on retail pricing.
   Use cost data to model how changes in supplier costs could affect pricing.
6. Estimate future supplier reliability based on current trends.
   Analyze current reliability data to predict future supplier performance.
7. Develop a model to select the best supplier based on various criteria.
   Create a decision model considering cost, quality, lead time, and reliability.
8. Simulate the impact of a new supplier entering the market.
   Model how a new supplier with different metrics might affect the supplier landscape.

Examples of Predictive Analytics in Supplier and Sale Data