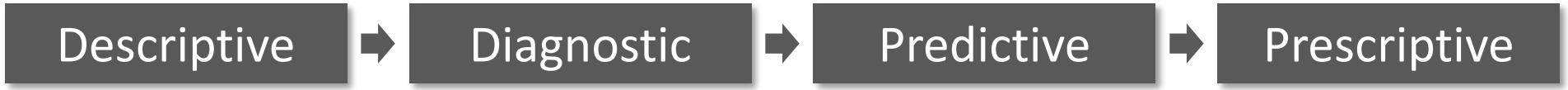
The Four Stages of Data Analytics

1. **Descriptive Analytics: What happened?**
2. **Diagnostic Analytics: Why did it happen?**
3. **Predictive Analytics: What might happen in the future?**
4. **Prescriptive Analytics: What should we do about it?**



Importance of Data Analytics in HR

- Strategic decision-making with data-driven insights.
- Applications: employee turnover, enhancing engagement, optimizing hiring processes.
- Data analytics provides actionable insights beyond intuition.

Descriptive Analytics

Descriptive → Diagnostic → Predictive → Prescriptive

- Examination of data to answer "What happened?"
- Summarizing raw data with dashboards, graphs, and tables.
- **Examples**: Football match scorecard, stock market trends.
- HR Example: Dashboard with HR metrics (employee satisfaction, cost per hire).

Descriptive Analytics

Descriptive → Diagnostic → Predictive → Prescriptive

- Common tools: Microsoft Excel, Tableau, Power BI, Google Analytics.
- Application: Visualizing turnover rates by department, tenure, and reasons for leaving.

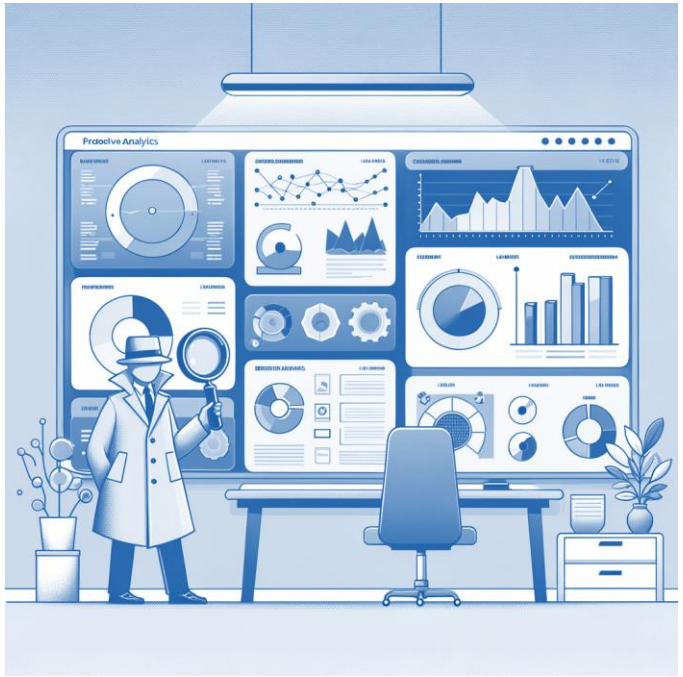
Diagnostic Analytics

Descriptive → **Diagnostic** → Predictive → Prescriptive

- Answering "Why did it happen?"
- Uncovering root causes and relationships.
- Importance in understanding variables influencing outcomes.
- Example: Analyzing the reasons behind stock market trends.

Diagnostic Analytics

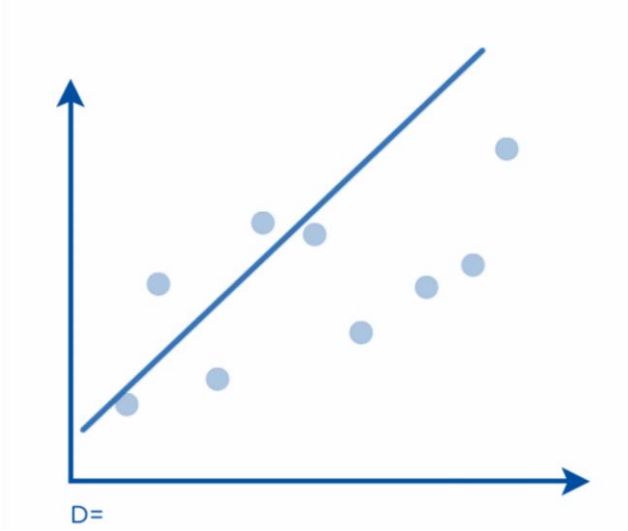
Descriptive → **Diagnostic** → Predictive → Prescriptive



- Common tools: Excel, SQL, Tableau.
- Application: Statistical analysis (correlation, regression) and pattern recognition.

Predictive Analytics

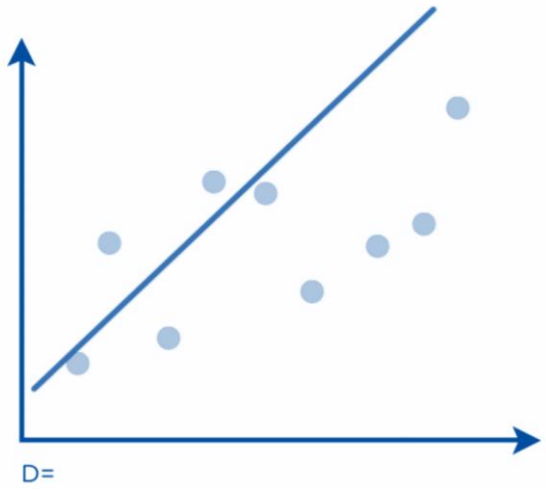
Descriptive → Diagnostic → **Predictive** → Prescriptive



- Using historical data and machine learning to forecast future outcomes.
- Applications: Predicting customer churn, employee attrition, sales forecasts.
- HR Example: Predicting staffing needs for workforce planning.

Predictive Analytics

Descriptive → Diagnostic → **Predictive** → Prescriptive



- Advanced tools: Python, R, SAS, Excel add-ons.
- Application: Time series forecasting, regression models, machine learning

Prescriptive Analytics

Descriptive → Diagnostic → Predictive → **Prescriptive**

- Suggesting actions based on predictions.
- Using machine learning, business rules, and algorithms.
- Providing recommendations for decision-making.
- Example: Advising on stock purchases.

Prescriptive Analytics

Descriptive → Diagnostic → Predictive → **Prescriptive**

- Application in HR: Planning hiring rounds, scheduling training, optimizing workforce utilization.
- Integration with predictive models for actionable insights.

People Analytics

Descriptive | Diagnostic | Predictive | Prescriptive

Recruitment and selection

Compensation management

Talent management, etc.

Metrics

Forecasting

Linear Regression

Clustering, etc.

People Analytics

HR Metrics

HR metrics, or human resources metrics, are key figures that help organizations track their human capital and measure how effective their human resources initiatives are

- Cost per hire
- Offer acceptance rate
- Hr to employee ratio

Choices of ratios may varies across industries/organizations

People Analytics

HR Metrics

1. Data-driven decision-making
2. Identifying trends and patterns
3. Benchmark for comparison
4. Foster accountability

HR Metrics

Staffing Metrics

1. **Time-to-Start**
   Average number of days it took to fill a position.

$$= \frac{\text{Total days elapsed from the date each filled position was available to the date each new person started in the position}}{\text{Number of positions filled}}$$

2. **Time-to-Productivity**
   Average number of days to satisfactory productivity

$$= \frac{\text{Total days elapsed from the date each filled position was available to the date each new person achieved satisfactory productivity}}{\text{Number of positions filled}}$$

3. **Accession Rate**
   Rate at which employees are joining the organization in a given time period.

$$= \frac{\text{Total Hire}}{\text{Regular headcount}}$$

4. **Turnover Rate**
   Rate at which employees are leaving the organization in a given time period.

$$= \frac{\text{Number of separations during the time period}}{\text{average actual number of employees during the time period}}$$

5. **Cost Per External/Internal Hire**
   Average cost incurred with an External/Internal hire.

$$= \frac{\text{Total costs related to all external/internal hires}}{\text{Number of external/internal hires}}$$

Other Staffing matrics

Offer Rate, Offer Decline Rate, Promotion Rate, Retention, Vacancy/Occupancy Rate, Vacancy Costs and Cost per Vacancy, Retirement Risk

HR Metrics

Sourcing Channel Metrics

1. Sourcing Channel Effectiveness (SE)

   Compares the number of applications a channel generates to the number of hires made from those applications. A lower SE indicates a more effective channel.

$$= \frac{Number\ of\ applictions\ generated\ through\ the\ channel}{Number\ of\ \ hired\ candidates\ through\ the\ channel}$$

2. Sourcing Channel Cost (CC)

   This metric measures the cost-effectiveness of a recruitment channel

$$= \frac{Amount\ of\ money\ spent\ on\ advertising\ on\ a\ channel}{Number\ of\ \ eligible\ candidates\ applied\ on\ that\ channel}$$

HR Metrics

Campus Recruitment Metrics

1. Campus Application Attraction Ratio (CAA))

Compares the number of applications a channel generates to the number of hires made from those applications. A lower SE indicates a more effective channel.

$$= \frac{Number\ of\ candidates\ applying}{Total\ number\ of\ eligible\ candidates}$$

2. Campus Selection Ratio (CS)

This metric measures the effectiveness of our campus recruitment

$$= \frac{Number\ of\ candidates\ hired}{Total\ candidates\ applying}$$

3. Offer Acceptance Ratio (OA):

This ratio indicates how many of our job offers are accepted

$$= \frac{Number\ of\ offer\ acceptances}{Total\ offers}$$

HR Metrics

Campus Recruitment Metrics

1. Average Retention Time (CRR)

   This metric shows the average duration that employees stay with the organization

$$= Average\ number\ of\ years\ of\ stay\ in\ the\ organization$$

2. Top Performer Index (TPI):

   his index measures the proportion of top performers among all hires.

$$= \frac{Number\ of\ top\ performers}{Total\ hired\ candidates}$$

KPIs

Metrics (KPIs)

1 **Clarity on the Objective**

KPIs

Metrics Example

Scenario:
Mr. Dan, the store manager, focused on the number of visitors to measure store performance. He is giving 100% off on tType equation here.he first order. However, despite high visitor numbers, financial expectations were not met due to discounts offered.

$$Total\ visitors: Number\ of\ visitors\ visited\ the\ store$$

Issue:
The KPI was misaligned with the actual performance objectives, leading to misleading conclusions.

$$Conversion\ rate = \frac{Number\ of\ visitors\ using\ services\ without\ discount}{Total\ number\ of\ visitors}$$

$$Total\ Daily\ Revenue:\ Measure\ the\ actual\ sales\ revenue\ generated\ daily$$

$$Lost\ rate = \frac{Total\ revenue\ lost\ in\ discount}{Total\ revenue}$$

Source: HR Analytics: Connecting Data and Theory

KPIs

Metrics

1	Clarity on the Objective
2	Decision Criteria
3	Data Collection
4	Analysis
3	Presentation of KPIs
4	Storytelling

Introduction to forecasting

Workforce Planning

Employee Turnover

Recruitment Forecasting

Absenteeism Prediction

Workforce Forecasting

What & Why

Sales forecasting is the process of estimating future staffing needs

Forecast can help in planning for future

Workforce Forecasting

What & Why

Time series forecast – Find patters in historical data

Time	Jan-22	Feb-22	Mar-22	Apr-22	May-22	Jun-22	Jul-22	Aug-22	Sep-22	Oct-22	Nov-22	Dec-22	Jan-23
Workers	6184	6105	6831	6430	6734	7118	7165	7436	8509	8213	9008	9309	??

Machine Learning models – Relationship with other factors

Business growth

Unemployment rate

Industry trends

Workforce

Performance Appraisals (PA) Metrics

Objectives of PA

1. **Evaluation of performance at a position:**

   Assess how well an employee is performing

2. **Development:**

   Identify areas for improvement and professional growth

3. **Documentation of records:**

   Maintain detailed records for future reference

Source: HR Analytics: Connecting Data and Theory

Performance Appraisals (PA) Metrics

PA Metrics

1. Department-Wise Promotions

$$\text{Promotion rate in a department} = \frac{\text{Number of promotions}}{\text{Total employees in the department}}$$

2. Lateral Movement Rate

$$\text{Lateral movement rate} = \frac{\text{Number of candidates provided lateral movement}}{\text{Total lateral requests}}$$

3. Average Promotion Rate

$$\text{Average promotion rate} = \frac{\text{Total sum of number of years of all employees served before being promoted}}{\text{Number of promotions in 10 years}}$$

4. Manager Instability Ratio

$$\text{Manager instability ratio} = \frac{\text{Number of employees worked under more than one manager}}{\text{Total employees in the department}}$$

Source: HR Analytics: Connecting Data and Theory

Performance Appraisals (PA) Metrics

PA Metrics

Monitor future leadership

1. Internal Succession Bandwidth

$$\text{Internal succession bandwidth} = \frac{\text{Number of positions where there is a ready internal candidate}}{\text{Total vacant positions in the succession planning}}$$

2. Fast-Track Employees

$$\text{Fast-track employees} = \frac{\text{Number of fast-track promotions}}{\text{Total promotions}}$$

3. Retention Rate of Fast-Track Employees:

$$\text{Retention rate of fast-track employees} = \frac{\text{Number of fast-track employees left in the organization at the year-end}}{\text{Total number of fast-track employees promoted in the last year}}$$

4. Top Management mix

$$\text{Top management mix} = \frac{\text{Number of senior-level positions filled with internal promotion}}{\text{Total number of senior-level positions}}$$

Performance Appraisals (PA) Metrics

PA Metrics

PA process monitoring

1. **Performance Appraisal Fill Rate**

$$\text{Performance appraisal fill rate} = \frac{\text{Number of employees filling PA form without help}}{\text{Total employees}}$$

2. **Appraiser Literacy Rate on PA**

$$\text{Appraiser literacy rate} = \frac{\text{Total number of trained appraisers}}{\text{Total number of appraisers}}$$

Compensation Management Metrics

Compensation Metrics

1. **Average Salary Position-Wise**

$$\text{Average salary} = \frac{\text{Sum total of the salary of all employees at a job position}}{\text{Total number of employees in that job position}}$$

2. **Within-Time Settlement**

$$\text{Percentage of within-time settlement cases} = \frac{\text{Total number of cases settled in time}}{\text{Total number of cases settled}}$$

3. **Salary Benchmarking**

Compensation Management Metrics

Compensation Metrics

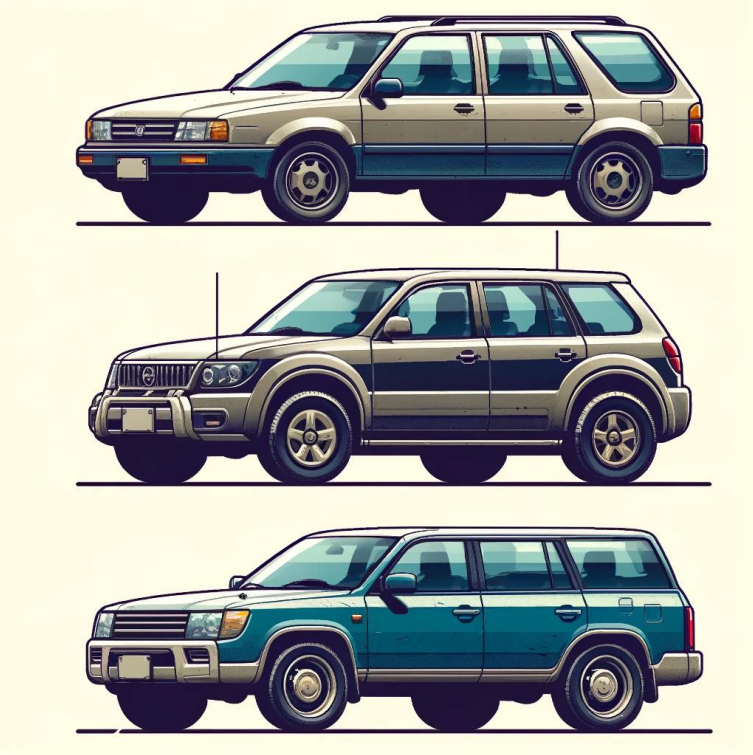
1. **Variance in Allocation**

$$\text{Percentage} = \frac{\text{Actual spending on rewards and incentives}}{\text{Total budgeted amount}}$$

2. **Within-Time Settlement**

$$\text{Percentage of within-time settlement cases} = \frac{\text{Total number of cases settled in time}}{\text{Total number of cases settled}}$$

Compensation Management



Regression

- Regression is all about predicting a continuous output variable based on one or more input variables
- Predict the price of a used car from factors like the car's age, mileage, and brand

Help John

John

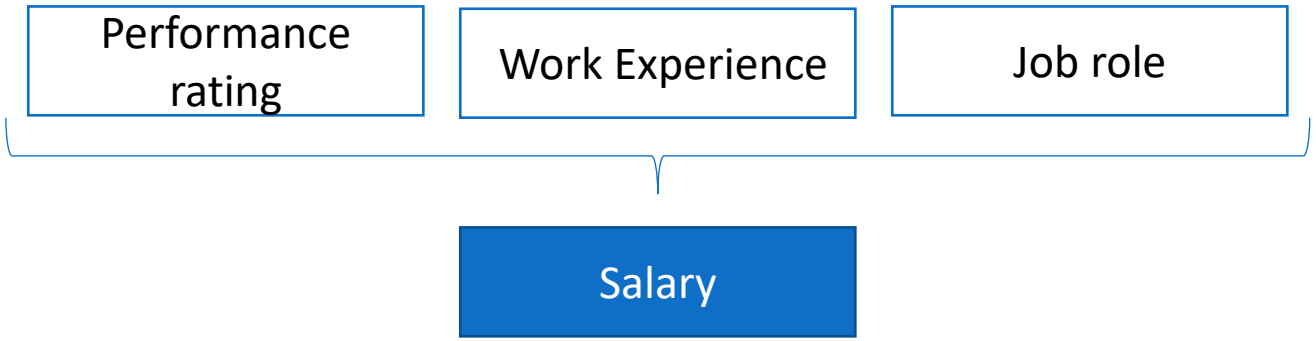
Data Analytics Manager at STA Corp

John, a data analyst at STA Corp, is working on an important project to examine salary disparities within the company. John aims to understand how different factors like job title, experience, education level, remote work status, city type, and performance ratings influence employee salaries. His goal is to identify any unfair differences in pay that might indicate bias or inequity.

Linear Regression

What & Why

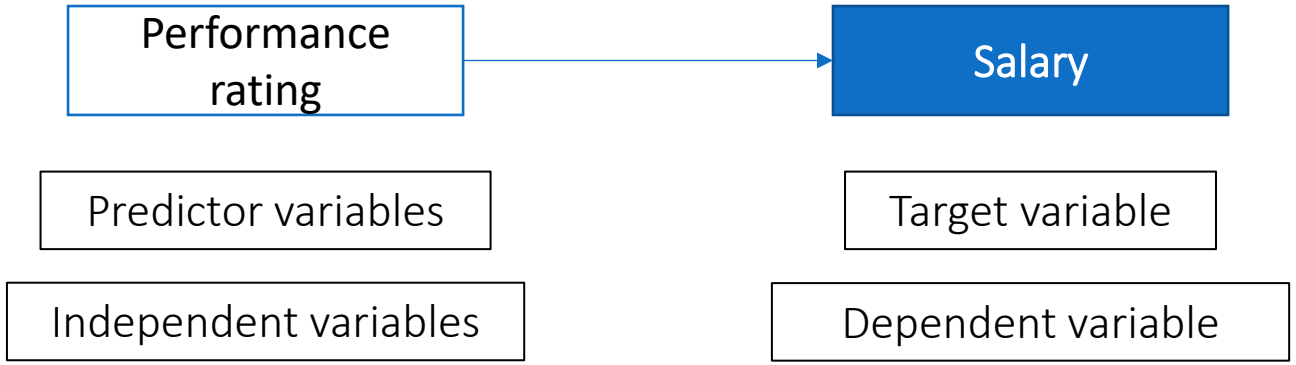
Find relationship between salary and other factors



If the performance rating increases by 1 point, what happens to sales?

Linear Regression

How



Linear relationship (single predictor variable)

$$y = ax + b$$

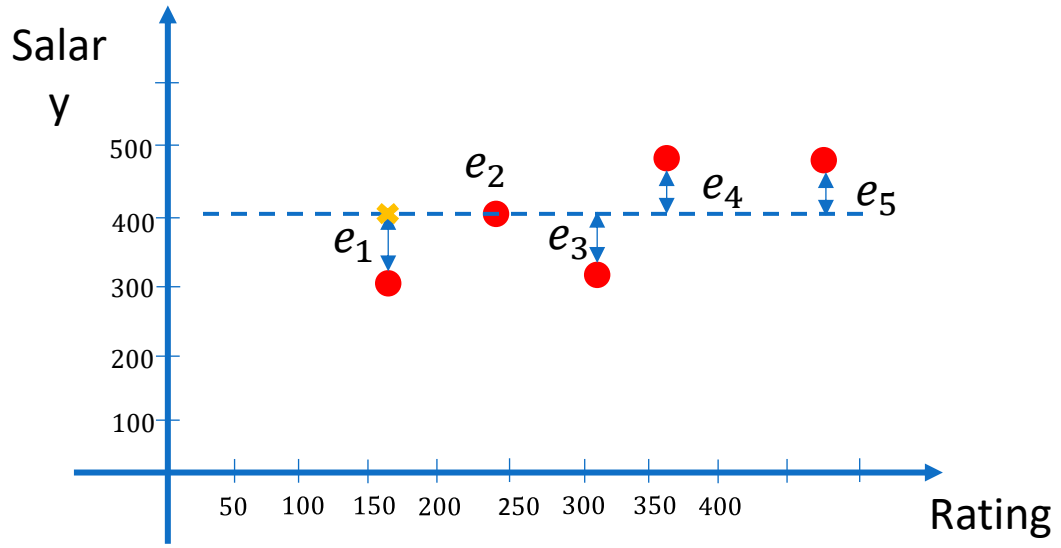
$$Salary = a \times (Performance\ rating) + b$$

Multiple predictors –

$$y = a_1x_1 + a_2x_2 + a_3x_3 + \cdots + a_nx_n + b$$

Linear Regression

How



$$Salary = a \times (Performance\ rating) + b$$

$e_1 + e_2 + e_3$…. Will not work

$$Minimize\ [e_1^2 + e_2^2 + e_3^2 + e_4^2 + e_5^2]$$

- This is called **minimizing sum of squared errors**

Linear Regression

Result

	A	B	C	D	E	F	G	H	I
1	SUMMARY OUTPUT								
2									
3	*Regression Statistics*								
4	Multiple R	0.97012							
5	R Square	0.941133							
6	Adjusted R Squa	0.926416							
7	Standard Error	515.7254							
8	Observations	11							
9									
10	ANOVA								
11		*df*	*SS*	*MS*	*F*	*Significance F*			
12	Regression	2	34017672.79	17008836	63.94955	1.20086E-05			
13	Residual	8	2127781.758	265972.7					
14	Total	10	36145454.55						
15									
16		*Coefficients*	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
17	Intercept	-1642.04	1151.801633	-1.42563	0.191804	-4298.10333	1014.015	-4298.1033	1014.01533
18									
19	Rating	8.133162	2.624163278	3.099335	0.014681	2.081830676	14.18449	2.08183068	14.1844934
20									

Coefficients – the a's and b's of this equation

$$y = a_1x_1 + a_2x_2 + a_3x_3 + \cdots + a_nx_n + b$$

Use these to create the final equation

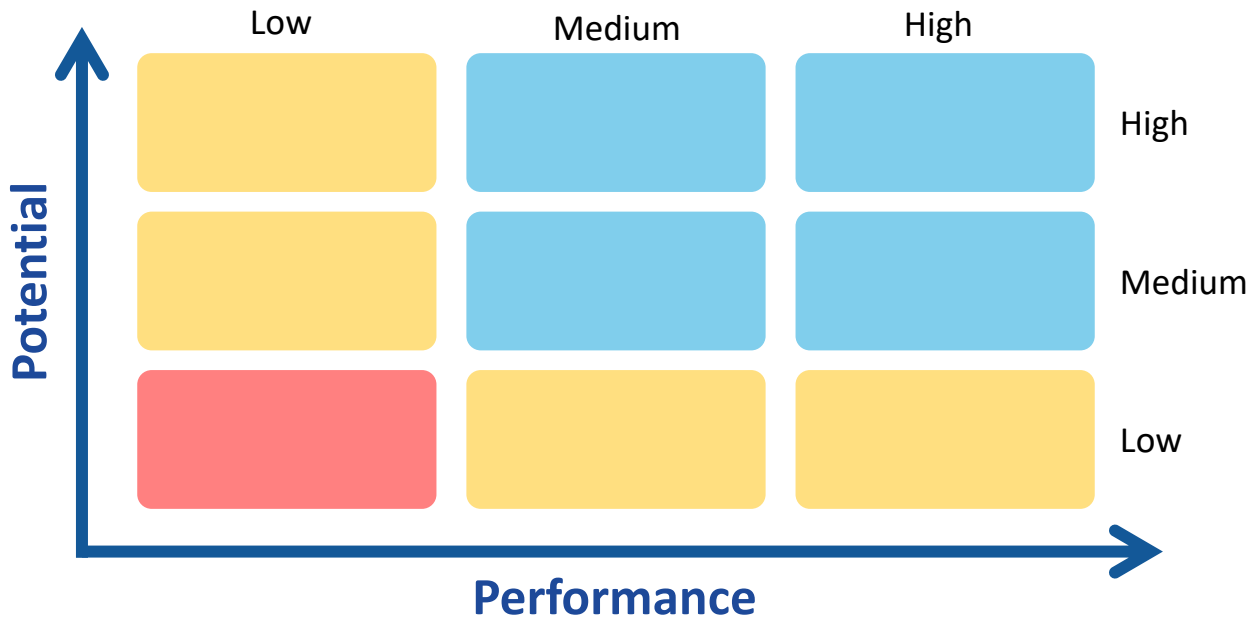
For eg. - $Salary = 8 \times (rating) - 250$

Talent Management Analytics

9-Box Grid

What is 9-Box grid model

9-box grid is a powerful talent management tool used to evaluate employees based on two dimensions: performance and potential

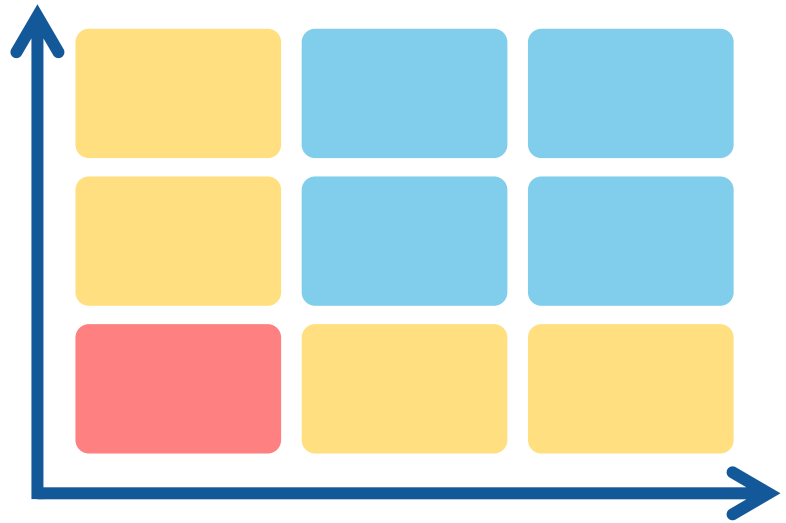


Talent Management Analytics

9-Box Grid

Advantages

1. Simplicity
2. Comprehensive Overview
3. Targeted Development
4. Enhanced Reviews



Talent Management Analytics

Performance Criteria

1. Low Performance

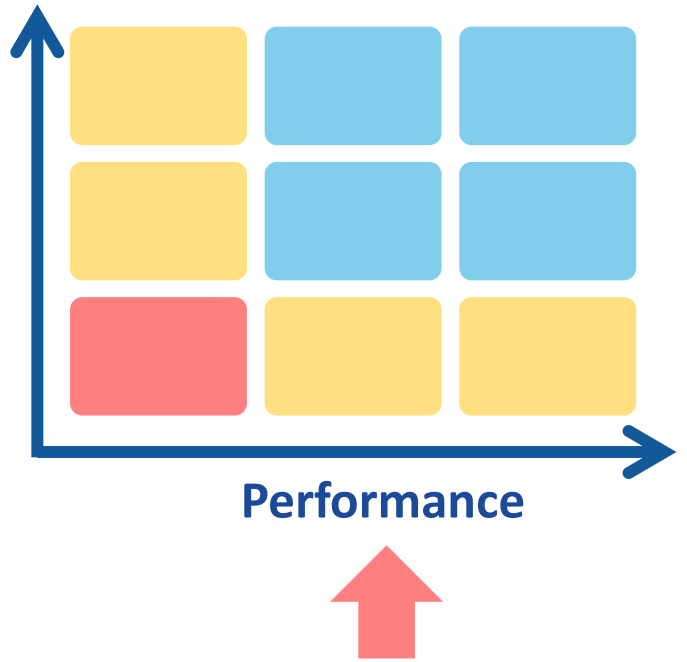
   Employees in this category consistently fail to meet job requirements and targets.

2. Moderate Performance

   Employees in this category meet some but not all job requirements.

3. High Performance

   Employees in this category consistently exceed job expectations and targets.



Talent Management Analytics

Potential Criteria

1. Low Potential

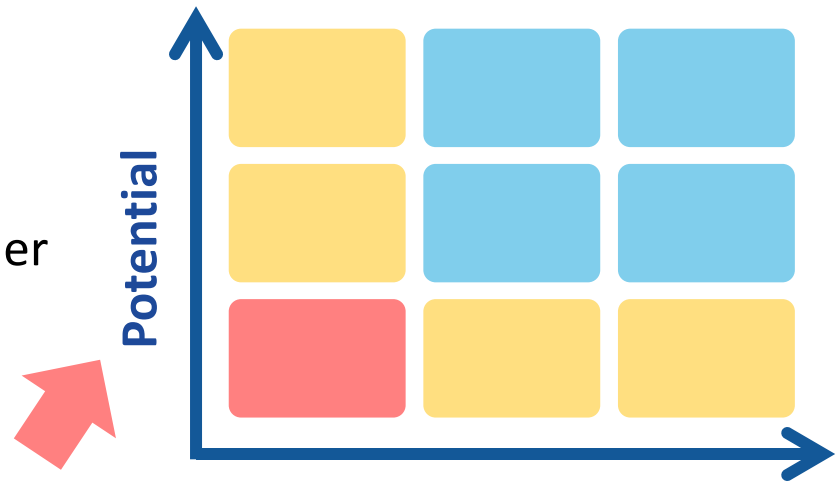
   Employees working at full capacity with no expected improvement.

2. Moderate Potential

   Employees capable of developing further within their current roles in terms of performance, skills, or expertise

3. High Potential

   Exceeding expectations, showing leadership qualities

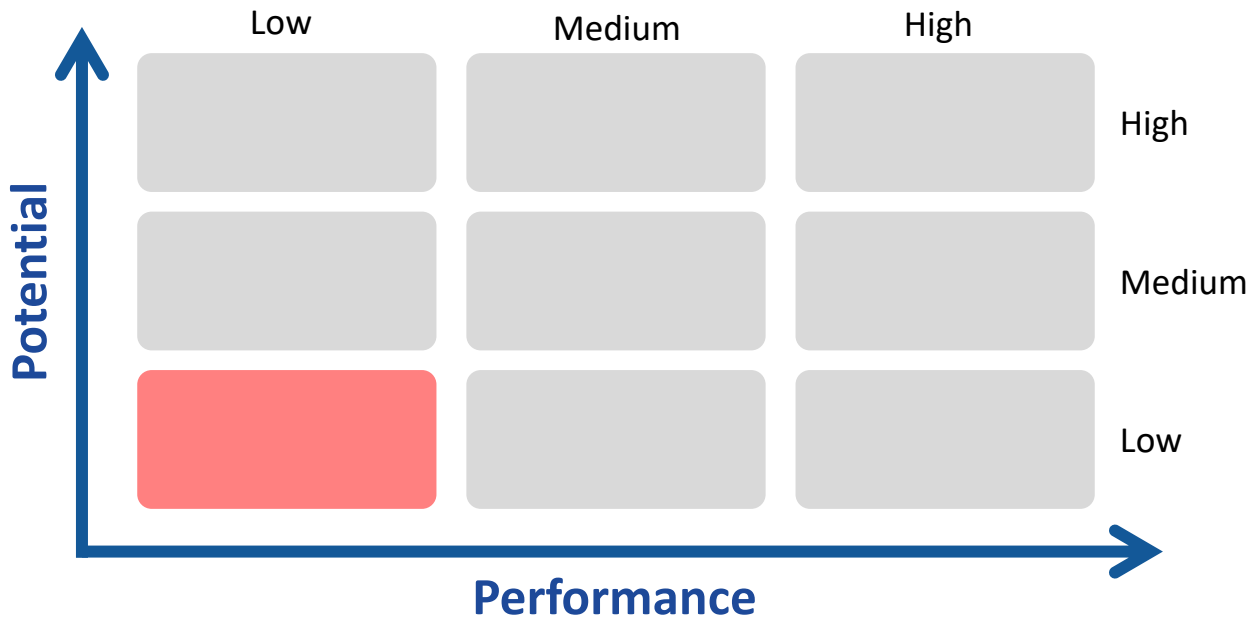


Talent Management Analytics

1. Observe/ Terminate

Low Performers with Low Potential

- Identify Roadblocks
- Reassign
- Exit Plan

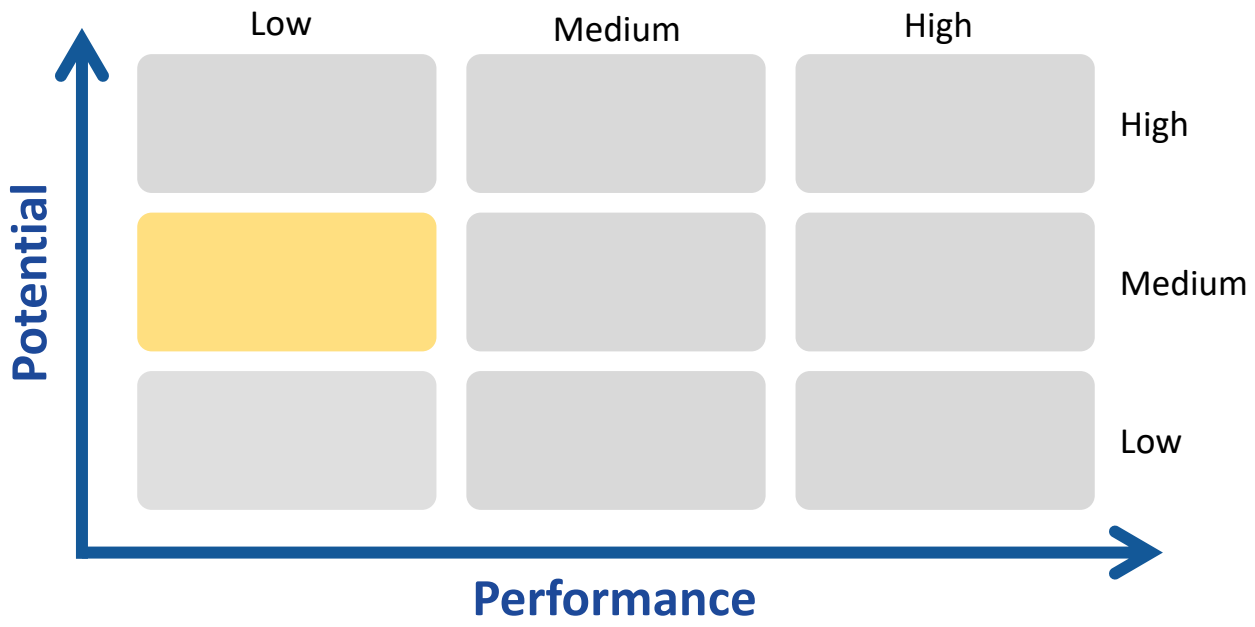


Talent Management Analytics

2. Observe Dilemmas

Low Performers with Medium Potential

- Performance Improvement Plan
- Monthly Check-ins
- Exit Plan

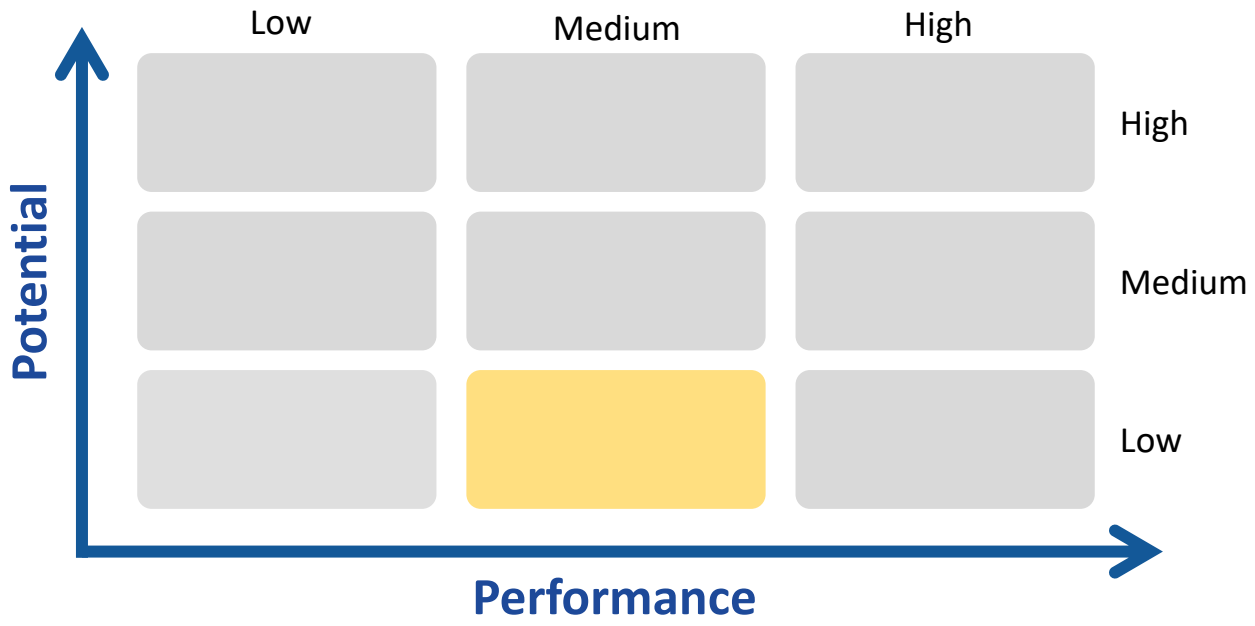


Talent Management Analytics

3. Observe Effective

Medium Performers with Low Potential

- Performance Improvement Plan
- Monthly Check-ins
- Exit Plan

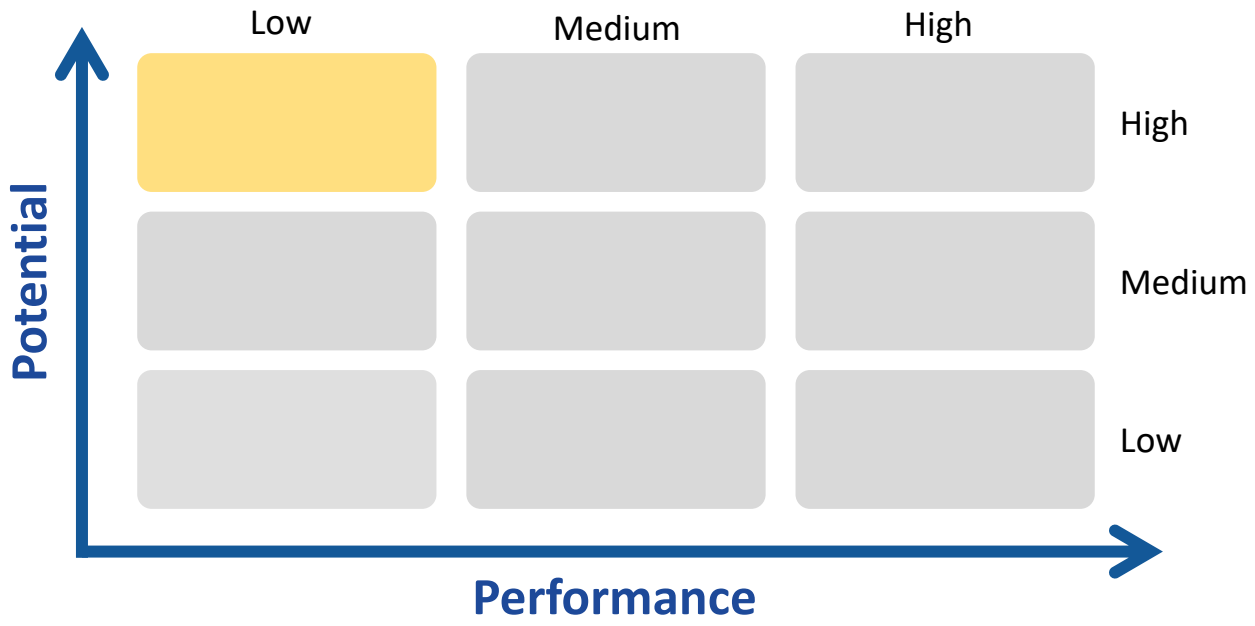


Talent Management Analytics

4. Develop professionals

Low Performers with High Potential

- Monitor Performance
- Communicate Expectations
- Exit Plan

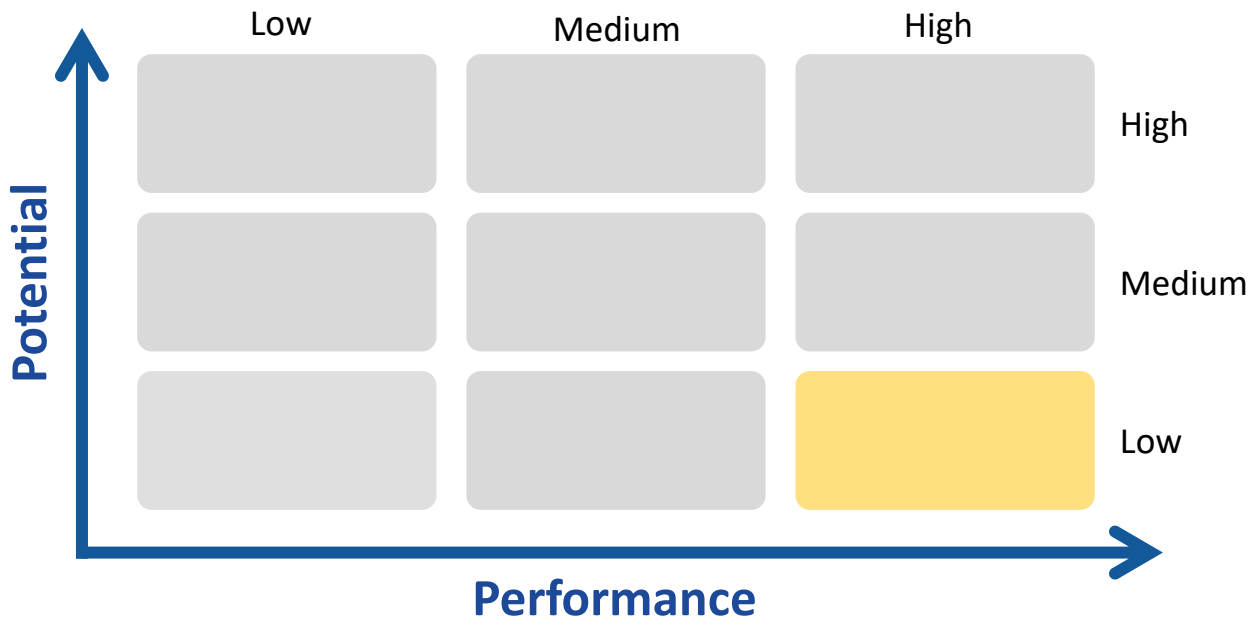


Talent Management Analytics

5. Trusted professionals

High Performers with Low Potential

- Keep Them Happy
- Future-Proof Career

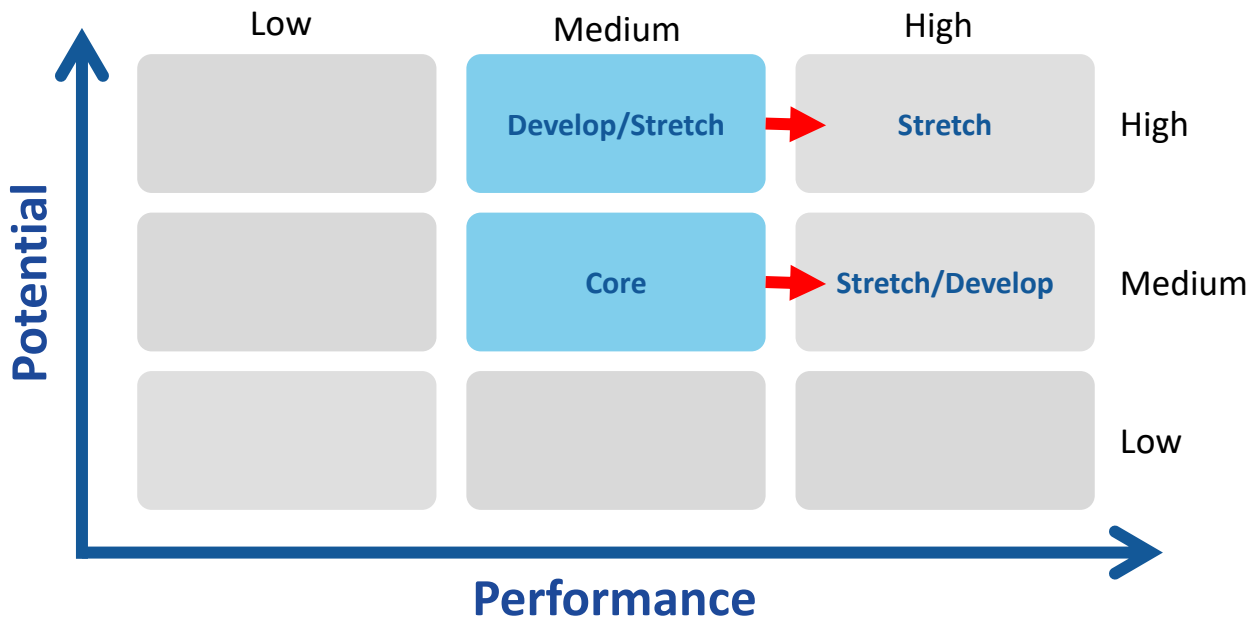


Talent Management Analytics

6-7. Develop/Stretch Core

Moderate to High Performance and Potential

- Set Clear Role Expectations
- Provide Job Exposure and Training
- Provide Coaching

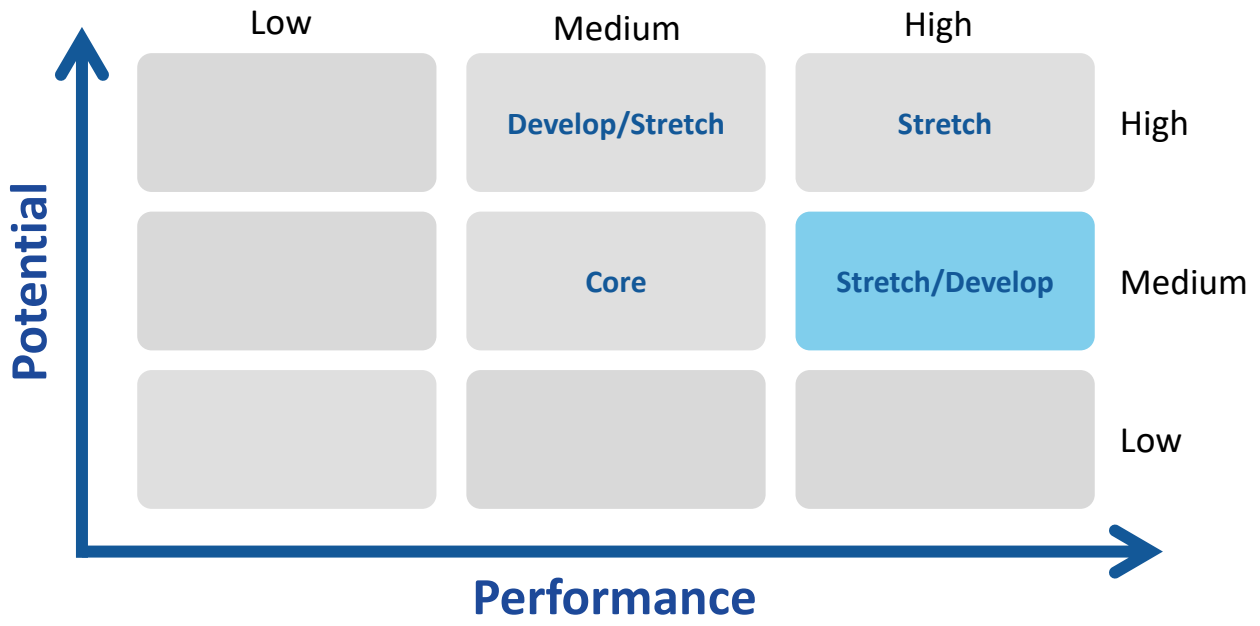


Talent Management Analytics

8. Stretch/Develop

Moderate to High Performance and Potential

- Motivate your high performers
- Foster growth potential
- Offer challenging assignments or job rotation

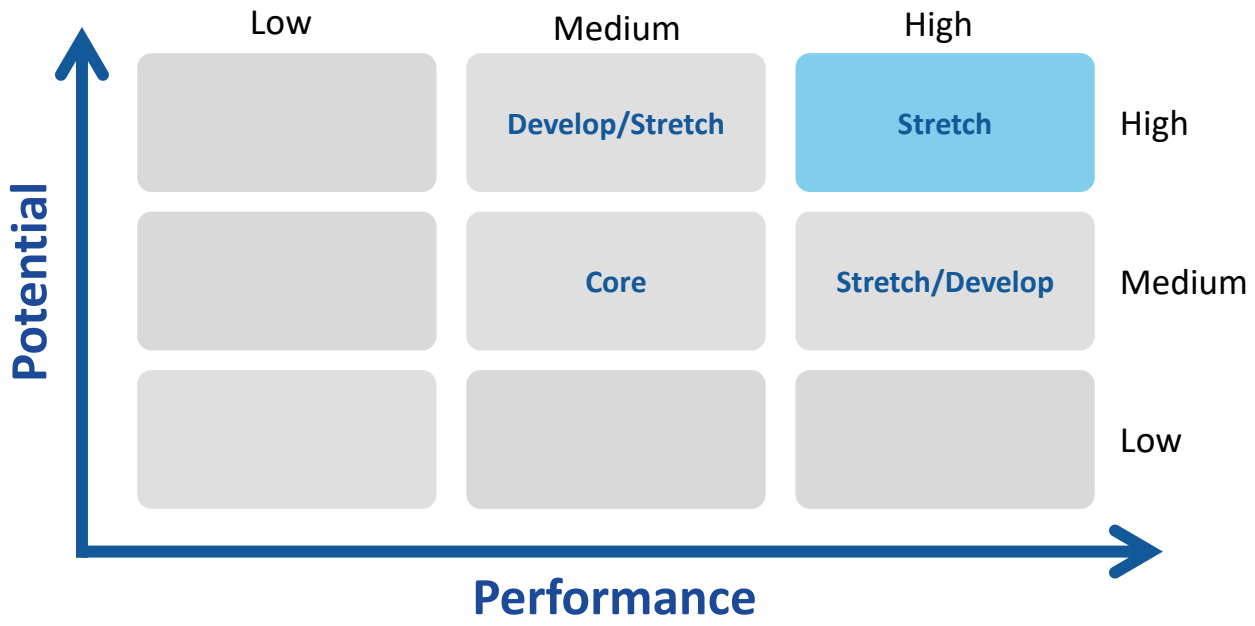


Talent Management Analytics

9. Stretch

High Performers with High Potential

- Offer Challenging Assignments
- Regular Check-ins and Appreciation
- Mentorship and Networking



Talent Management Metrics

Talent Turnover or Churn Rate

Talent turnover, or churn rate, is the attrition of employees from an organization

1. **Voluntary Turnover**: When employees leave on their own accord, often for better career opportunities elsewhere.
2. **Involuntary Turnover**: When employees leave due to retirement, layoffs, or termination.

$$\text{Turnover Rate} = \frac{\text{Number of employees left in a particular quarter}}{\text{Number of employees that we started in the last quarter}}$$

Classification

Classification

- Classification is used when we want to predict a categorical output
- A bank might want to determine whether a loan applicant is high-risk or low-risk

Retention

What

Extend employee tenure by preventing churn

Hiring cost, Training time, lowers productivity etc

Overall churn rate or attrition rate is the key KPI for HR leadership

Time Series forecasting to predict overall churn

ML models to predict probability of churn

Logistic regression can help predict probability of churn

Pitch alternate options/ offers to the employees to retain the employees

Logistic Regression

What

Uses historical data to train, same as linear regression

Linear Regression – predicts continuous variable

Logistic Regression – predicts categorical variable

Output of logistic regression corresponds to probability of belonging to one of the categories

Logistic Regression

Example

Working hours (per week)	Commute distance (Km)	Tenure (Months)	Churned?
40	15	34	0
45	25	6	1
42	23	10	0
…	…	…	…

Use this data to train our logistic regression model

Working hours (per week)	Commute distance (Km)	Tenure (Months)	Churned?
50	20	12	0.75
35	10	29	0.32

Logistic Regression

Confusion Matrix

	Predicted Positive	Predicted Negative
Actual Positive	True Positive	False Negative
Actual Negative	False Positive	True Negative

True Positive - the model correctly predicted the churn.
False Positive - the model incorrectly predicted the churn.
True Negative - the model correctly predicted the retention.
False Negative - the model incorrectly predicted the retention.

Accuracy = (True positive + True Negative)/ Total cases

	Predicted Positive	Predicted Negative
Actual Positive	63	27
Actual Negative	38	122

Accuracy = (63+122)/250 = 185/250 = 74 %

Help Tim

Tim

HR Analytics Manager at STA Telecom

TA Solutions is an IT company that provides software development and IT support services to various clients

Tim has been tasked with reducing employee attrition. He believes that he can use a logistic regression model to predict which employees are likely to leave the company. After identifying these employees, he thinks that STA Solutions can improve retention by addressing their concerns and offering incentives.

Training and Development Metrics

Training and Development Metrics

1. Training Man Days

$$\text{Training Man Days} = \frac{\text{Total number of days}}{\text{Total employees}}$$

2. Position-Wise Employees Trained

$$\text{Position-Wise Employees Trained} = \frac{\text{Employees trained in a position}}{\text{Total employees in a position}}$$

3. Diversity of Training

$$\text{Diversity of Training} = \frac{\text{Number of new training introduced}}{\text{Total training in a year}}$$

4. Type of Trainer

$$\text{Type of Trainer} = \frac{\text{Training by internal trainer}}{\text{Total trainers}}$$

Training and Development Metrics

Training and Development Metrics

1. **E-Learning Type**

$$\text{E-Learning Type} = \frac{\text{Total e-learning trainings}}{\text{Total trainings}}$$

2. **Participants Satisfaction**

$$\text{Participants Satisfaction} = \frac{\text{Satisfaction score}}{\text{Number of employees}}$$

3. **Absenteeism in Training**

$$\text{Absenteeism in Training} = \frac{\text{Number of employees absent}}{\text{Total training participants}}$$

4. **ROI of the Training**

$$\text{ROI of the Training} = \text{Financial Gain} - \text{Training Cost}$$

Training and Development Metrics

Training and Development Metrics

1. Competency Rate

   Degree to which employees in key positions have the competencies necessary to achieve their performance objectives.

   $$= \frac{(\#\ of\ incumbents\ with\ competency\ ratings\ of\ Acceptable\ or\ better}{\#\ of\ incumbents\ who\ have\ received\ competency\ assessments)}$$

2. Readiness

   Reflects how ready the organization is from a human capital perspective to execute on strategy and achieve key goals and objectives

   $$= Occupancy\ Rate\ x\ Competency\ Rate\ X\ 100$$

3. Training Spend Rates

   Relative importance of spend on training vs other operating and human capital activities.

   $$= (Training\ spend/\ Total\ Human\ Capital\ Spend)\ x\ 100$$

4. Training Headcount investment

   Amount of investment in trainings for each employee

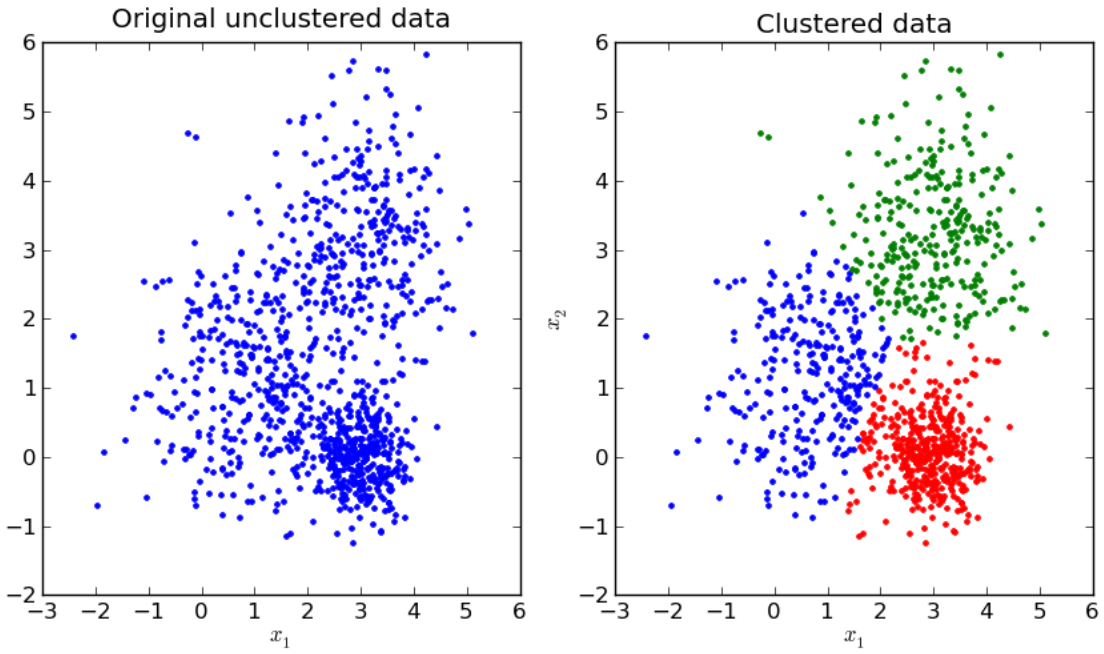
   $$= \frac{Total\ training\ cost}{Regular\ headcount}$$

Clustering

Clustering

- Clustering is a technique used to group similar data points together based on their characteristics
- For instance, a marketing team might want to segment customers into different groups based on purchasing behavior, demographics, and preferences

Introduction to Clustering

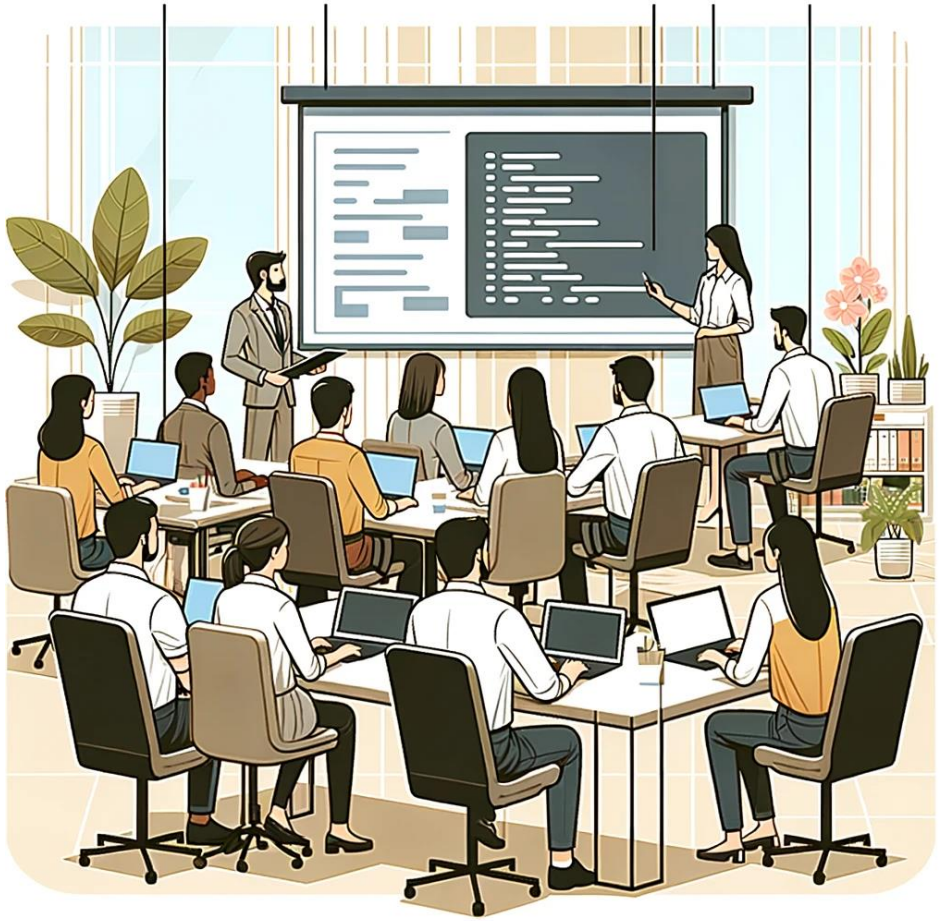


- **Definition**: Clustering is a technique used in data analytics to group objects in such a way that objects in the same group (or cluster) are more similar to each other than to those in other groups.
- **Unsupervised Learning**: Unlike forecasting, regression, and classification, clustering is an unsupervised learning method, meaning it does not rely on predefined labels or outcomes. Instead, it discovers the natural groupings within the data.

HR Example

- **Scenario:** Imagine you are an HR professional tasked with dividing a group of 300 employees into three distinct groups for training and development.
- **Data Collection:** Each of the 300 employees has been assessed on three criteria: technical skills, communication skills, and presentation skills. The scores range from 1 to 10 for each skill.

Example



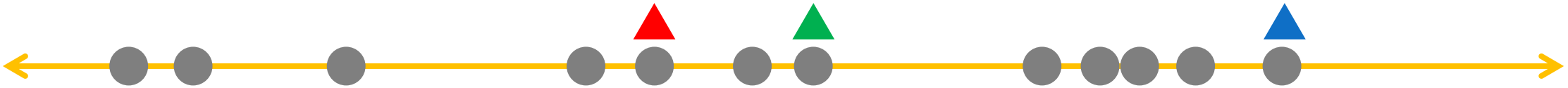
Cluster Analysis:

1. Cluster 1: Employees with high technical skills but lower communication and presentation skills.
2. Cluster 2: Employees with moderate scores across all three skills.
3. Cluster 3: Employees with high communication and presentation skills but lower technical skills.

K Mean Clustering

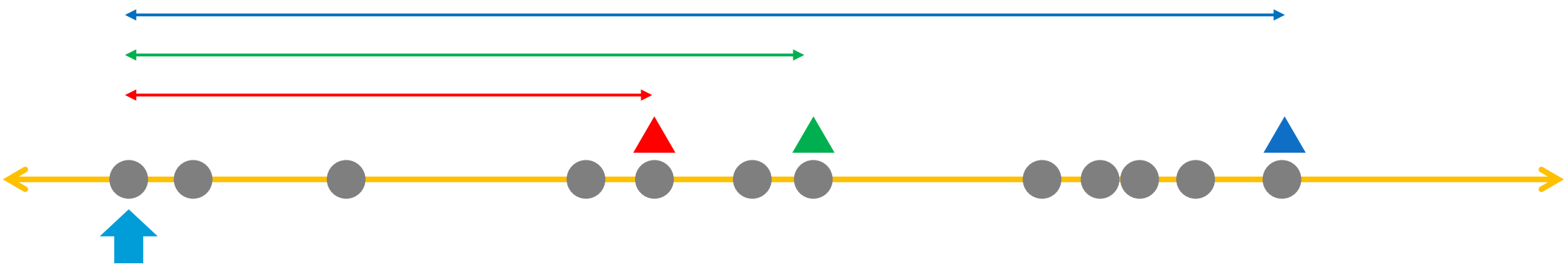
K Mean Clustering

Randomly choose 3 points (centroids)



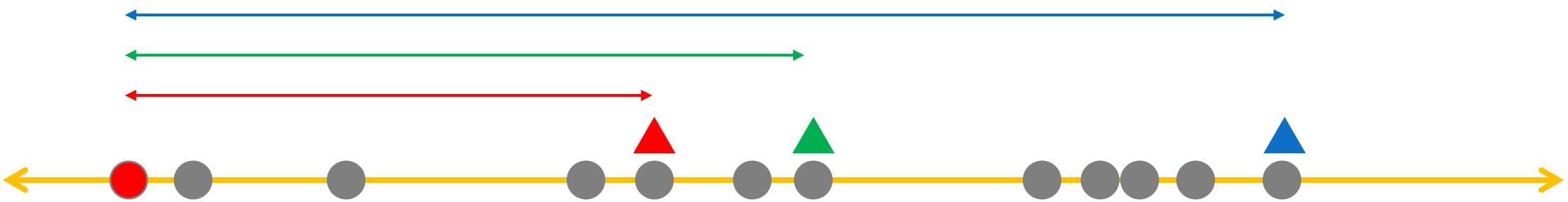
K Mean Clustering

Measure distance of a point with 3 centroids



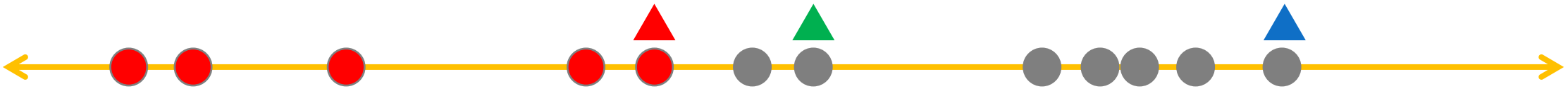
K Mean Clustering

Measure distance of a point with 3 centroids



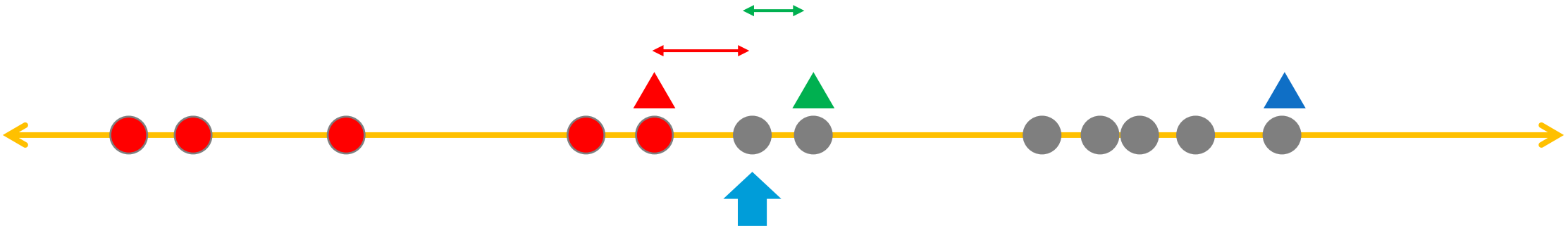
K Mean Clustering

Measure distance of a point with 3 centroids



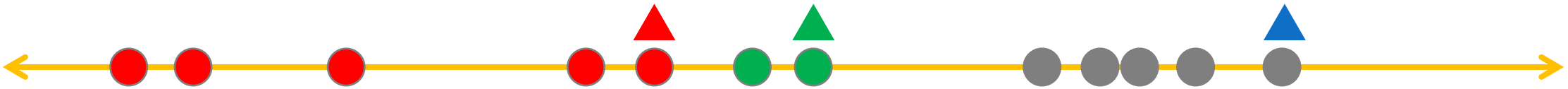
K Mean Clustering

Measure distance of a point with 3 centroids



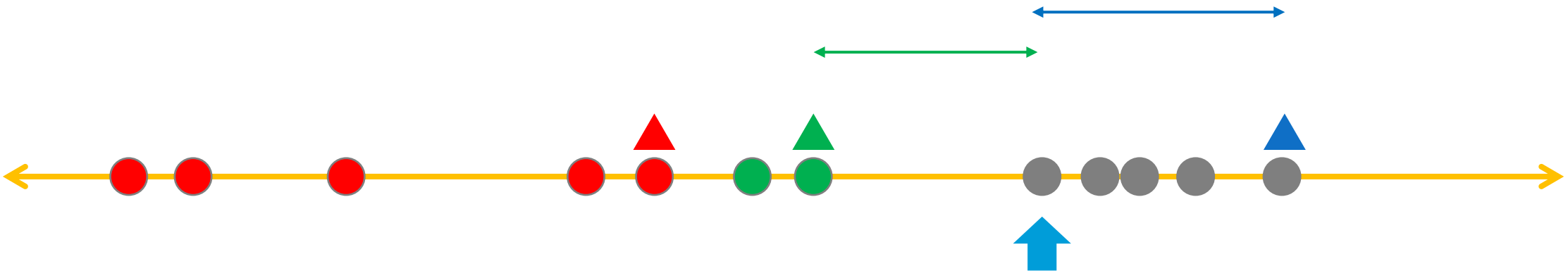
K Mean Clustering

Measure distance of a point with 3 centroids



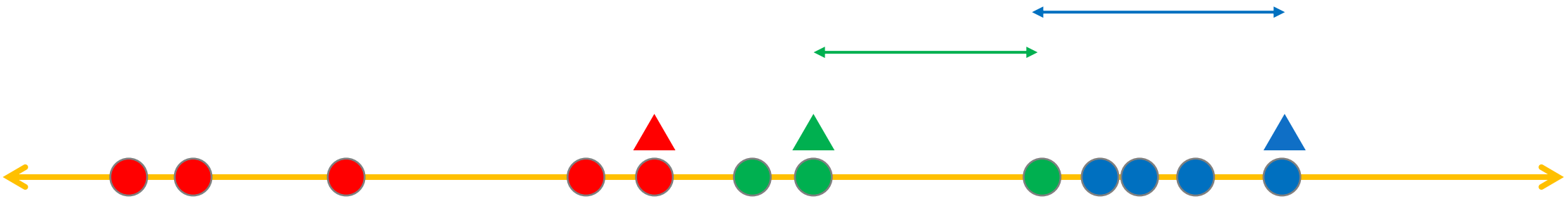
K Mean Clustering

Measure distance of a point with 3 centroids



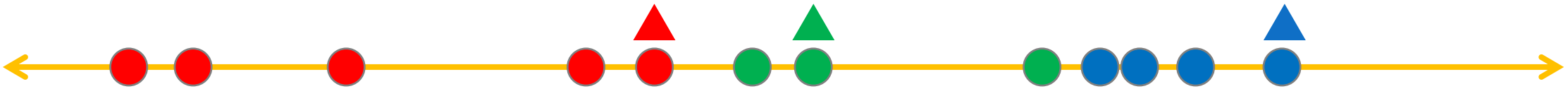
K Mean Clustering

Measure distance of a point with 3 centroids



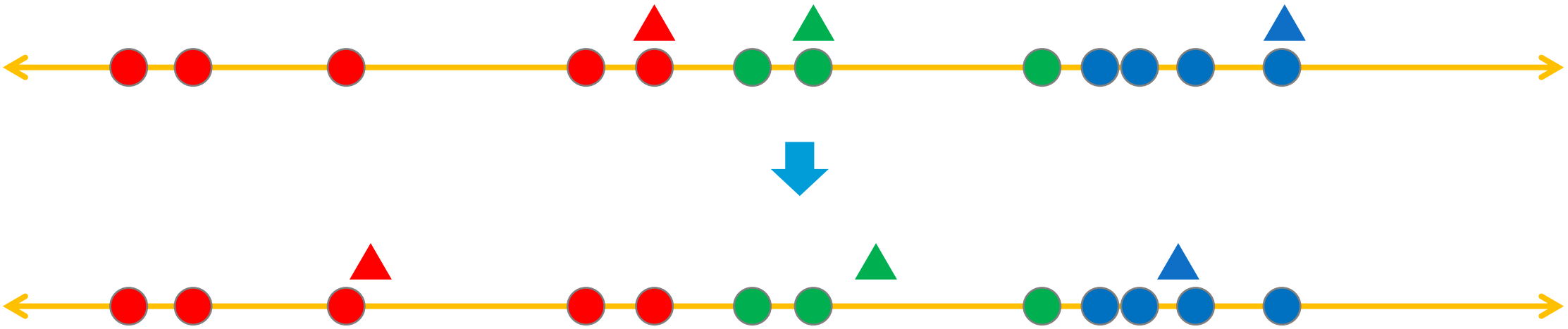
K Mean Clustering

End of 1st cycle



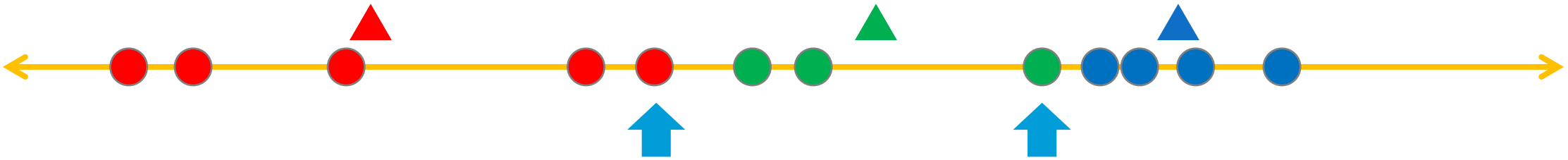
K Mean Clustering

Take the average of each cluster and shift the centroid to that point



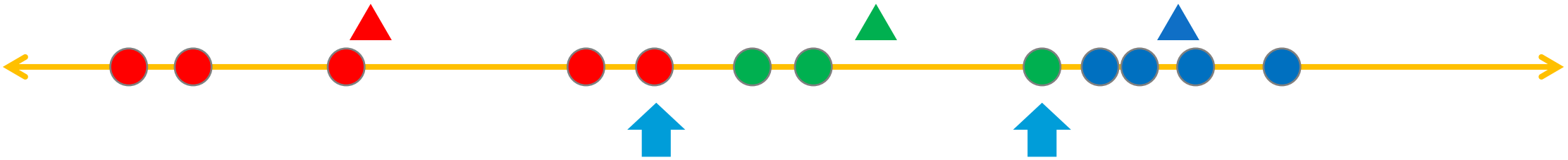
K Mean Clustering

Measure distance of a point with 3 new centroids and reassign



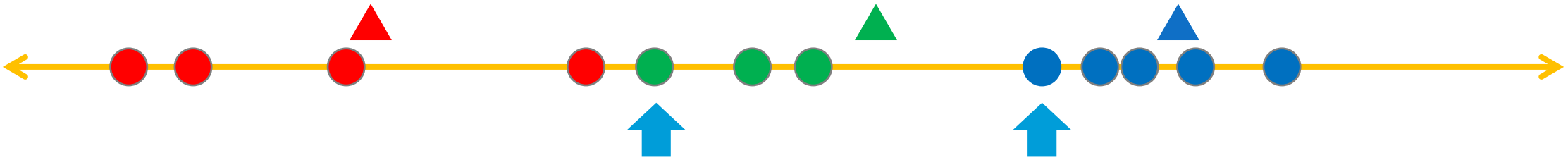
K Mean Clustering

Measure distance of a point with 3 new centroids and reassign



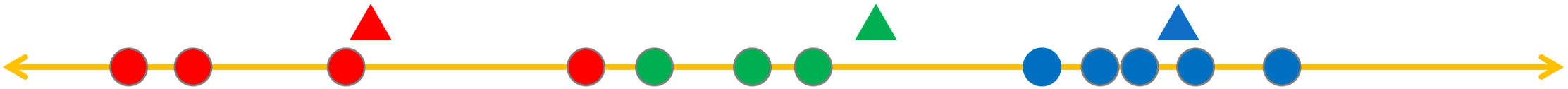
K Mean Clustering

Measure distance of a point with 3 new centroids and reassign



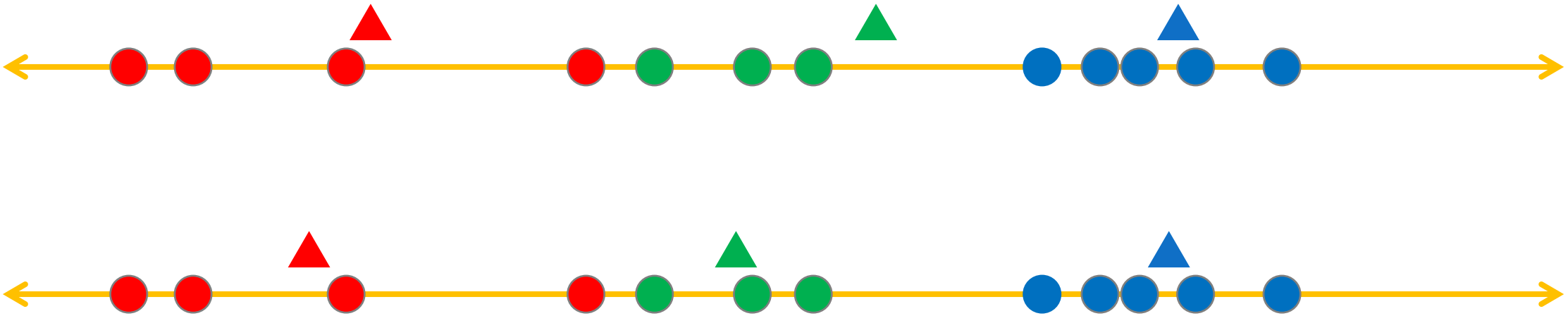
K Mean Clustering

End of cycle 2
Now again, take the average of each cluster and shift the centroid to that point



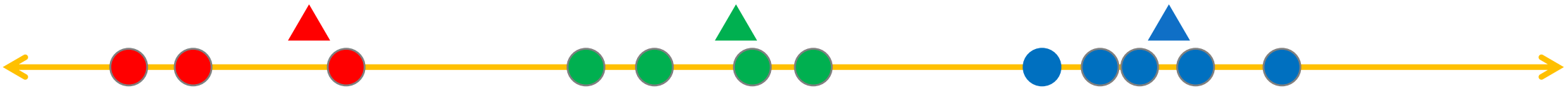
K Mean Clustering

Now again, take the average of each cluster and shift the centroid to that point



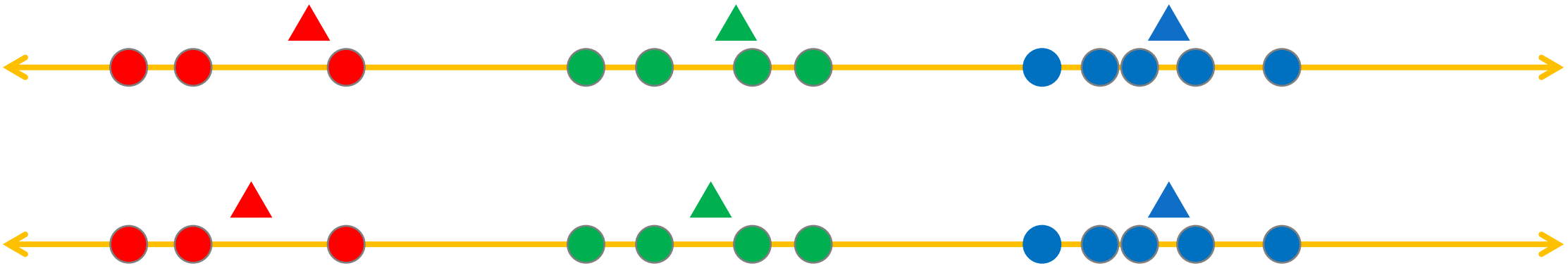
K Mean Clustering

Reassign points



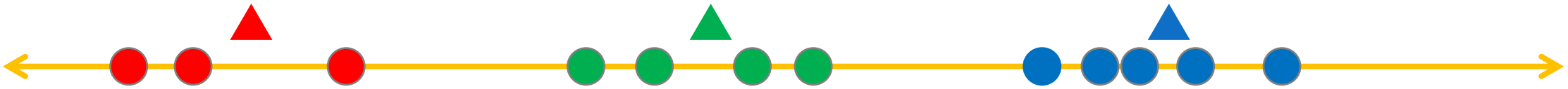
K Mean Clustering

End of cycle 3. Now again, take the average of each cluster and shift the centroid to that point



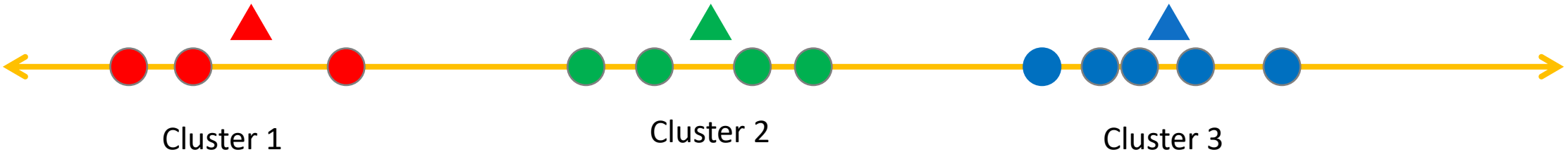
K Mean Clustering

Measure distance of a point with 3 new centroids and reassign



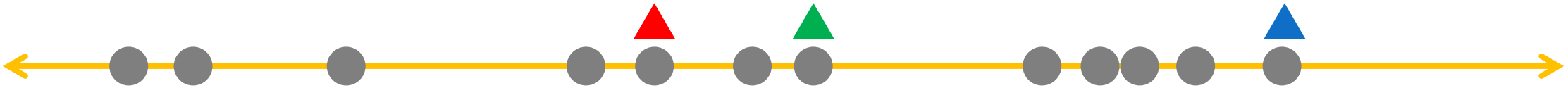
K Mean Clustering

If no changes, STOP



K Mean Clustering

Randomly choose 3 points (centroids)



Help Jim

Jim

HR Analytics Manager at STA Telecom

Jim has been tasked with designing effective training programs for employees. He believes that he can use a K-means clustering model to group employees based on their skill levels in technical, communication, and presentation skills. After identifying these groups, he thinks that STA Solutions can enhance training and development by tailoring programs to meet the specific needs of each cluster.