This is a graded discussion: 6 points possible

due Apr 24 at 4:29pm

Customer Segmentation in the Airline Industry [Interactive Presentation and Discussion 4.1]

22 68

- (https://classroom.emeritus.org/courses/9054/modules/items/1506954)
- (https://classroom.emeritus.org/courses/9054/modules/items/1506956)
- (https://classroom.emeritus.org/courses/9054/modules/items/1506957)
- (https://classroom.emeritus.org/courses/9054/modules/items/1506959)
- (https://classroom.emeritus.org/courses/9054/modules/items/1506960)
- (https://classroom.emeritus.org/courses/9054/modules/items/1506961)
- (https://classroom.emeritus.org/courses/9054/modules/items/1506962)
- (https://classroom.emeritus.org/courses/9054/modules/items/1506964)
- (https://classroom.emeritus.org/courses/9054/modules/items/1506965)
- (https://classroom.emeritus.org/courses/9054/modules/items/1506966)
- (https://classroom.emeritus.org/courses/9054/modules/items/1506967)
- (https://classroom.emeritus.org/courses/9054/modules/items/1506969)
- (https://classroom.emeritus.org/courses/9054/modules/items/1506971)
- (https://classroom.emeritus.org/courses/9054/modules/items/1506972)
- (https://classroom.emeritus.org/courses/9054/modules/items/1506973)

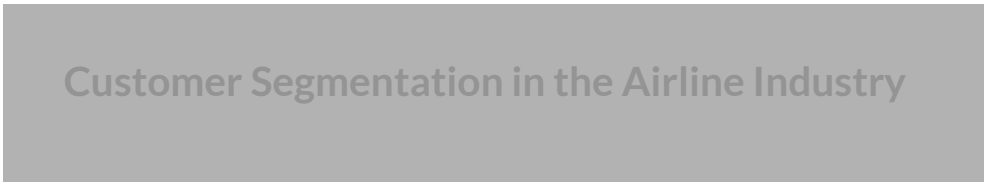
There is one interactive presentation and one discussion on this page. Please scroll down to view and complete them.

Interactive Presentation: Customer Segmentation

Live Support

in the Airline Industry (Data)

Next, you will go through the following interactive presentation on customer segmentation in the airline industry. In the presentation, Retsef introduces the case and the data you will be using to create clusters.



Resume

Restart

Discussion 4.1: Making Business Decisions Through Comparing Clusters [20 Minutes]

🎯 Learning Outcome Addressed:

- Make business interpretations of clusters obtained by algorithms.

This is a required discussion and will count toward course completion.

Given the use cases presented by Retsef in this module, consider the types of clustering that can be useful to your current industry. Can you identify specific areas in your organization that would

benefit from clustering?

Be sure to read the statements posted by your peers. Engage with them by responding with thoughtful comments and questions to deepen the discussion.

Suggested Time: 20 minutes

Rubric: Discussion 4.1

Criteria	Exceeds expectations	Meets expectations	Below expectations
Thoughtful and complete response to the question(s)	**4 pts** Fully responds to the question(s), post is supported by connections to the reading and real-life examples, and post makes additional connections to the field of data engineering with novel ideas, critical thinking, or extensive application of how to use the topic in future work.	**3 pts** Fully responds to the question(s), and post is supported by connections to the content or real-life examples.	**0 pts** Partially responds to the question(s), or connections to the content are missing or vague.
Engagement with the learning community	**2 pts** Posts thoughtful questions or novel ideas to multiple peers that generate new ideas and group discussion.	**1.5 pts** Asks questions or posts thoughtful responses to generate a single peer's response.	**0 pts** No responses to peers or posts minimal or vague responses to peers that do not motivate a response (e.g., "I agree.").

Reply

Diego Milanes (*He/Him*) (https://classroom.emeritus.org/courses/9054/users/228518)

Apr 17, 2024

Several programs at the university are intended to identify students at potential socio-economic risk. I think that clusterisation can play an important role in this process. A first unsupervised clusterisation algorithm can potentially identify different types of students based on the admission scores, notional test scores, and prior socio-economic information often requested when a student is admitted to a public university. This initial categorisation is a starting point and can be used as a first approach to an academic triage system and anomaly detection.
Afterwards, supervised learning (knowing an outcome tag), including academic scores during the first years of university, can be performed, adding information to the algorithm such as profiles of desertion, profiles of successful students, and profiles of students requesting processes at the wellness and mental health assistance offices. A deeper categorisation with this information helps in the early detection of at-risk students, and prevention alerts can be triggered on time.

Reply

Turki Alghusoon (https://classroom.emeritus.org/courses/9054/users/229165)

Apr 20, 2024

Hi Diego,

Interesting idea! I think this could significantly promote equity and improve the experience of the students. do you envision the algorithm to eventually feed into a prescriptive model where at-risk students could be automatically flagged for mental/physical health assistance by university professionals? could the model also be expanded to include financial risk clustering?

Best,

Turki

Reply

Diego Milanes (*He/Him*) (https://classroom.emeritus.org/courses/9054/users/228518)

Apr 22, 2024

Hi Turki

Indeed, that is the idea. To have a trained algorithm to flag them well in advance problems start to manifest.

I found financial risk to be more complex since one might need additional (sensitive) data but this is something to work on.

cheers

Diego

Reply

Manjari Vellanki (https://classroom.emeritus.org/courses/9054/users/231480)

Apr 17, 2024

In Clinical firm, recently I have gone through an article that an analysis has been performed on all clinical trials data to automatically identify and cluster clinical trials with similar eligibility features by extracting features from eligibility criteria text of all clinical trials and constructed a trial-feature matrix. And then calculate the pair wise similarities for all clinical trials based on their eligibility features. For all trials, by selecting one trial as the center each time, then identified trials whose similarities to the central trial were greater than or equal to a predefined threshold and constructed center-based clusters. Then unique trial sets have been identified with distinctive trial membership compositions from center-based clusters by disregarding their structural information. This approach can be potentially useful for investigating knowledge reuse patterns in clinical trial eligibility criteria designs and for improving clinical trial recruitment.

Reply

Roy Nunez (https://classroom.emeritus.org/courses/9054/users/229552)

Apr 18, 2024

Similar to Retsef's airline example, my team in the banking industry can segment customers based on their monetary transactions ie. credit card and debit cards using transaction

frequency, transaction amounts, and we can categorize the products and services purchased, etc and include their net worths as data points.

We can use the clustering to separate clusters such as high-value customers, bargain seekers, less frequent shoppers, loyal customers.

This clustering will help us with more personalized marketing by improving out understanding of the behaviors and needs of each customer group. Some examples of the personalized marketing include credit card offers that offer more reward points for a specific category like dining, travel, or higher spending habits.

Another idea for clustering is using clustering to segment our customers in another four groups, like a wealthy group, a young generation, an average group, and conservative customers. Similarly to the previous market segments we can use monetary transactions and balances.

We can target the younger generation with digital banking tools, starting credit cards, education on financial literacy. We can target a wealthier group with premium services and investment products along with some exclusive offers. The average group can be targeted with standard credit services and products. The conservative customers can be targeted with savings products or even investment products.

Ultimately these data driven insights when done right will increase the effectiveness of the marketing efforts, help us understand our customers better and generate better matches with products and services that will serve their needs.

Reply (1 like)

(http

Javier Di (https://classroom.emeritus.org/courses/9054/users/226884)

Apr 19, 2024

Great example Roy and interested if you can do a clustering analysis by wealth/transaction value use Vs Fees? And if you get the most fees from the wealthier customers to direct most of the marketing dollars to them, where you're getting the most money? (not sure if this is the case but asking)

Thanks

Reply

Roy Nunez (https://classroom.emeritus.org/courses/9054/users/229552)

Apr 20, 2024

Hi Javier,

This is a pretty good idea to consider. We currently havent been analyzing the fees, partly because we are a bit limited on this information. Only fees I can recall are interest charges on credit cards. May be some wealth advisor fees we can get from another team. Will keep this suggestion in mind and bring this up soon. I appreciate your suggestion!

Reply

Priscilla Annor-Gyamfi (https://classroom.emeritus.org/courses/9054/users/226376)

Apr 23, 2024

Hi Roy,

Great way of using clustering to group customers based on their characteristics, usage of their credit cards to identify trends/purchase pattern which can help target specific potential needs of customers and meet them through insightful marketing strategies and product/service recommendations.

Reply

Roy Nunez (https://classroom.emeritus.org/courses/9054/users/229552)

Apr 25, 2024

Thanks Priscilla! I hope we get the targeting and insights right :)

Reply

Mariana Flores (https://classroom.emeritus.org/courses/9054/users/237198)

Apr 23, 2024

Hi Roy, so nice to meet you. Great post, cluster analysis on behavioral observations in the banking industry to segment customers for marketing and to create personalized experiences is such a business impactful manner to implement cluster analysis. I’m with

you in that data-driven insights increase the effectiveness of marketing efforts by helping to better understand customers.

Cluster analysis and the real-world application is truly fascinating - thank you for sharing.

Reply

Roy Nunez (https://classroom.emeritus.org/courses/9054/users/229552)

Apr 25, 2024

Thanks Marianna. Yes, I agree and I look forward to applying our learned methods on real world applications.

Reply

Yossr Hammad (https://classroom.emeritus.org/courses/9054/users/229118)

Apr 18, 2024

While i was listening to restef, i was thinking that clustering maybe useful for Employees, we can applied for employment segmentation for example; we can segment employees based on some factors that would help the business.

We can segment them based on job roles, or performance, or skillset. it would give the business insight of the high performing teams, and assist those needs some learnings to get better. it also could help reduce the redundant roles that has same day to day tasks.

Also Clustering can be applied on Salaries; we can group employees with similar salary/salary grade , benefits, bonus and etc... that can help the business to benchmark salaries and ensure equitable compensation ...

Another example is apply clustering on budget, it will help if we segment teams based on their budget utilization or profitability.. it will help in deciding budget allocation decisions and cost optimization!

Reply (1 like)

Javier Di (https://classroom.emeritus.org/courses/9054/users/226884)

Apr 19, 2024

Yossr, I really liked the idea of classifying the employees with similar title/grade and I would include their reviews Vs salary and see how that clusters. That would not only tell about equitable compensation but give a sense for political favors/factors and other distortions that are likely affecting compensation and make the process more equitable. Great idea. Thanks, Javier

Reply (2 likes)

(http

Diego Milanes (*He/Him*) (https://classroom.emeritus.org/courses/9054/users/228518)

Apr 22, 2024

Hi Yossr

Great post! Have you considered how clustering techniques can be effectively utilized to optimise the various aspects of employee management?

thanks!

Diego

Reply

(http

Yossr Hammad (https://classroom.emeritus.org/courses/9054/users/229118)

Apr 23, 2024

Interesting point. if you can mention how or what are the factors that we can use to segment employees ,that would be great.

Thank you

Reply

(https:/

Javier Di (https://classroom.emeritus.org/courses/9054/users/226884)

Apr 19, 2024

I can think about some very useful applications of Clustering to the investment field in which I work.

2 specific examples on the investing side, where we want to obtain attractive returns and find ways to consistently achieve those returns are mapping the following on the X & Y axis to identify the relationship and useful methods to obtain high returns:

1) Returns on Capital of the stock/business (t0) Vs Future Returns (t1): generally businesses/stocks that have a high return on capital will have a future expected high return as the business itself is very profitable and obtaining very high returns on the capital that it is reinvesting and utilizing. We would expect the data to cluster around high returns on capital/high future returns and same for low returns on capital/ low returns;

2) Purchase Valuation (t0) Vs Future Returns (t1): businesses/stocks that are purchased at attractive valuations will have higher expected future returns as you have more room for the valuation to expand and increase as the sentiment or business results get better. This can develop a profitable and interesting investment strategy. We would expect the data to cluster around low valuations (t0)/ high future returns (t1) and high valuations (t0)/ low future returns (t1);

3) Application for fundraising: the ability of an investment firm to raise money is a generally mis-understood and opaque area and nobody really understands why some funds can raise money and others don't. A helpful application for clustering would be to analyze past month returns Vs Fundraising responses and see if the data groups on years of past high returns and getting more interest/ responses for fundraising after the strong returns (as commonly said "investors chase returns")

Reply

(http

Manjari Vellanki (https://classroom.emeritus.org/courses/9054/users/231480)

Apr 22, 2024

Interesting example :)

Reply

(http

Roy Nunez (https://classroom.emeritus.org/courses/9054/users/229552)

Apr 23, 2024

Hi Javier,

Thank you for your post.

Will you be also considering other factors that influence the outcomes of the complex dynamics of the financial markets? Interested to what factors if so.

Also I am still trying to understand how clustering will help in this scenario? I am understanding that there is already assumptions that high return on capital will have a future high return and conversely low return on capital will yield lower returns and similarly there are assumptions for Purchase Valuation.

Reply

Haitham Farag (https://classroom.emeritus.org/courses/9054/users/233864)

Apr 19, 2024

Industry Background

An international non-governmental organization (NGO) with a significant portion of income (donation) coming from individual donors. The key consideration is the return on investment of each marketing campaign while also competing with other NGOs.

Individual donor main parameters and characteristics (cluster dimensions)

- Age
- Gender
- Marital Status
- Donation amount in the past 12 months
- Number of Donations
- Media response rate (e.g. online ads, text, post phone calls etc)
- Response history to similar campaigns (war in Ukraine, war in Gaza)
- Time of year (e.g. seasonality, more likely to respond during a certain period Christmas or Easter)

Of course income and employment status would have been more impactful but my organization does not have access to it.

Clustering Recommendation

1. Divide the donation cause feature into type (e.g. war, earthquake) and the target country/ies for better behaviour clustering (of the individual donors). For example, the war in Ukraine and

the war in Gaza did not have the same traction among the clusters thought to be the most responsive. Accordingly, both the cluster parameters and their respective weight will need to be adjusted based on each campaign detail.

2. Assess how different units of measurement for some parameters (e.g. age bins/ different age groups) may affect the number and size of the clusters.

Limitation

Missing key data for some donors (e.g. age) .

Edited by Haitham Farag (https://classroom.emeritus.org/courses/9054/users/233864) on Apr 21 at 7:32am

Reply (1 like)

(http

Turki Alghusoon (https://classroom.emeritus.org/courses/9054/users/229165)

Apr 20, 2024

Hi Haitham,

Great example and the breakdown is very useful! Do you think normalizing some of the parameters (e.g., age, and donation amounts) could help stabilize the clustering?

Best,

Turki

Reply (1 like)

(http

Haitham Farag (https://classroom.emeritus.org/courses/9054/users/233864)

Apr 23, 2024

Very valuable recommendation.

Thanks Turki

Reply

(http

Gustavo Santana (https://classroom.emeritus.org/courses/9054/users/120927)

May 1, 2024

Hello Haitham, a great example of an interesting subject!

Thinking in this situation, it may be interesting to use clustering the number of donations based on Age/Gender/Marital Status and Response History to detect the donor's cluster with more probability to help in further campaigns.

Reply

Haitham Farag (https://classroom.emeritus.org/courses/9054/users/233864)

May 26, 2024

The number of Donations as an outcome is an interesting dimension Thanks for sharing Gustavo.

Reply

Roman Jazmin (https://classroom.emeritus.org/courses/9054/users/225803)

Apr 19, 2024

I worked for a company that focuses on winning government contracts in bidding auctions. Based on this week's lesson, what would be very beneficial for my industry is detecting big rigging using cluster analysis.

Bid rigging tends to be an on-going problem for both Governmental and public authorities in an environment where fair play competitiveness should thrive and not eliminated. As stated in 1 article, "bidders concluded anticompetitive agreements to increase the bidding price and get better contracts from the public authorities."

Substantial problems in the awarding process includes excessive state intervention, discrimination in awarding contracts, and favoritism for local contractors. Cartels are formed by bidders to increase the auction price and get contracts at the auctioneers' costs.

Based on the nature of the bidding that is observed, we can apply behavioral screening as performed during the auction and we can observe certain aspects that may suggest agreements between the participants in the auctions. Thus, certain behavioral patterns of bidders can be captured, related either to the way of bidding or to the subsequent development of contracts, which may be the result of pre-arranged agreements.

To analyze the existence of a cartel on a given market, adequate quantitative method is represented by statistical analysis. It's necessary to use more variables, not just starting. and bid awarding price. such variables are sales capacities, transportation, experience, etc. The analytical method to detect bid-rigging was the cluster method when we do not have any information above the existence of a cartel.

When performing a cluster analysis, we divide the data into two groups. The first cluster included those auctions with a high ratio, when the sale price was close to the starting price, with the second cluster that includes auctions with a low ratio, sale price was significantly lower than the starting price, likely to be the result of competitive behavior.

Considering the selection of the two clusters, we expected the first cluster, for which the ratio was high, to have a normal distribution, while the second cluster, for which the ratio was

low, had an asymmetric distribution. This result would confirm the hypothesis that the cluster for which the ratio was low corresponds to the situation of competition, while the second cluster would correspond with the situation where, among the participants, there was a collusive agreement, because anticompetitive agreements maintain an extremely low bidding level close to the starting price.

Thus, an important area that my firm can benefit from clustering is its finances when bidding for a contract. We will be able to predict or determine how much cost we can expect from each project as well as the overall profit from a contract we have been awarded.

Reply

(http

Dawn Prewett (https://classroom.emeritus.org/courses/9054/users/233112)

Apr 21, 2024

This is a very interesting use case that you posit here and your methodology seems quite sound - especially since you were able to confirm your hypothesis. Once the data is in hand, however, it can be difficult to take action. What type of action can be taken in a situation where it is believed that collusion is occurring? In most cases, I would assume that the output that you obtained from your data cannot be directly used as proof, but only as a suggestion that something isn't right.

Reply

(http

Haitham Farag (https://classroom.emeritus.org/courses/9054/users/233864)

May 26, 2024

Thanks, Roman for the detailed business case. I found Dawn's question on how this can be used to be interesting also.

Reply

Turki Alghusoon (https://classroom.emeritus.org/courses/9054/users/229165)

Apr 20, 2024

Within my current organization, one immediate opportunity to apply clustering would be in project management. Most projects within my unit are expected to be completed within 150 calendar days (elapsed time), and to require less than 1,200 staff hours (cost). However, actual performance of projects varies significantly across elapsed time and cost.

We could utilize clustering to identify the performance cluster for all projects within the portfolio. By plotting project against cost and duration, we can identify the clusters of high performance, low performance, and the ones in between. In this specific scenario, projects' actual performance is measured as a percentage of the set targets. Since both metrics are percentages, there will be no need for normalization in this instance.

Once we have identified those clusters, we can then dive deep into each cluster as follows:

- **High-performing cluste**r: identify common best practices and behaviors within the clusters, and standardized them across the units, to promote project excellence.
- **In-between cluster:** understand the factors preventing this cluster from performing at the higher level and take steps to upskill the teams involved to move them up the performance curve.
- **Low-performing cluster:** identify common pitfalls and obstacles, develop mitigation strategies, and train staff on those strategies to minimize re-occurrence in the future.

Reply (1 like)

Shahrod Hemassi (*He/Him*) (https://classroom.emeritus.org/courses/9054/users/224267)

Apr 21, 2024

Hi Turki. I like your post as I work with a project management software that is focused on cost control. I like your idea, but I think it would be interesting to do the analysis of performance at a lower level. Instead of looking at the Project-level, maybe try looking at the data related to the people on the projects. By clustering the people that are involved in a project based on different metrics such as role, skillset, location, and performance, you

may find better indicators of which projects will perform well. This could give more insight into why some projects are performing better than others and identify more detailed areas to invest into in order to improve overall project performance. Interesting topic. Thanks for sharing.

Reply

Turki Alghusoon (https://classroom.emeritus.org/courses/9054/users/229165)

Apr 23, 2024

Hi Shahrod,

Thank you for the suggestion. I totally agree: analyzing the performance of the staff would be the natural next step to really get to the root causes for variation in performance.

Cheers,

Turki

Reply

Koffi Henri Charles Koffi (https://classroom.emeritus.org/courses/9054/users/208039)

Apr 29, 2024

hi Turki , nice one , a similar approach is used in my company to group employee base on their performance and technologies stack to assign them work that need to be accomplish in a period of two weeks

Reply

Priscilla Annor-Gyamfi (https://classroom.emeritus.org/courses/9054/users/226376)

Apr 23, 2024

Great submission Turki.

However, I would like to know whether cost and duration for the projects are the only two factors your organization considers to score performance? If that's the case I think you could consider normalization by introducing some equally important key factors such as

the expertise needed and the impact of project on a scale of 1-10 which will help make the clustering more effective and aid in well informed decision making.

Reply

Turki Alghusoon (https://classroom.emeritus.org/courses/9054/users/229165)

Apr 23, 2024

Hi Priscilla,

Great suggestions. Including data on the impact of projects would provide important context to the analysis, as the most impactful projects are sometimes the longest to complete.

Best,

Turki

Reply

Mariana Flores (https://classroom.emeritus.org/courses/9054/users/237198)

Apr 23, 2024

Hi Turki, so nice to meet you. Great post, cluster analysis in project management to identify project performance is such an impactful manner to implement cluster analysis. I'm with you in that learning from each cluster could provide insightful takeaways for future application.

Cluster analysis and the real-world application is truly fascinating - thank you for sharing.

Reply

Dawn Prewett (https://classroom.emeritus.org/courses/9054/users/233112)

Apr 21, 2024

I work with data obtained via audits of various processes and outputs over a number of different programs company-wide. Each audit can contain any number of findings and each finding is correlated to a code that represents the type of issue that was identified. Once a finding is created, the program is given time to respond, contain and correct. The finding itself is also given a risk rating, which is based on the type of issue that was identified. So, there is

a direct correlation between the code and the weight. Each stage of this process represents a different milestone that is marked by a due date that is based on a set of rules as programs are given x number of days to complete each stage.

A couple of interesting thoughts came to me while watching the videos and considering this data. It would be interesting to cluster the data by the finding code and the average time to each milestone. We can often tell on the outset whether the milestones for the corrective plan will be met or not simply based on the type of issue identified. Verifying this and identifying others, might help us proactively help the programs we're issuing findings to. Furthermore, those that take longer to correct may be inherently more risky because the problem is more complex or may linger longer after the finding is issued. It would also be interesting to see if similar findings cluster over time. Another aspect worth exploring is the temporal clustering of similar findings to identify recurring issues, which is currently a challenge due to the variability in how different auditors may categorize findings.

It would also be beneficial to cluster different finding types and even dates to completion against the auditor themselves. While our auditors are focused and incredibly professional; they still add a human element to the mix. Identifying whether certain findings are more likely to be reported by specific auditors could help pinpoint potential weaknesses or blind spots in our auditing process.

Reply

Ricardo Anaya (https://classroom.emeritus.org/courses/9054/users/228915)

Apr 22, 2024

time between issues, and yes human situation context and awarness is always good.

The human factor is laways there,. However, the more data, the more varaibles to automate and correlate, the human gets more simplified, not irrelevant, but more data driven.

Reply

Dawn Prewett (https://classroom.emeritus.org/courses/9054/users/233112)

Apr 22, 2024

That's a truly intriguing sentiment there: "the more data, the more variables to automate and correlate, the human gets more simplified". I think generalizing human behavior is easy when it is a group, but trying to figure out what a single person will do and it becomes difficult, which feels counterintuitive. However, maybe it shouldn't be

since more people equals more data at a faster rate and allows for greater generalizations. Thank you for the comment!

Reply

Shahrod Hemassi (*He/Him*) (https://classroom.emeritus.org/courses/9054/users/224267)

Apr 21, 2024

I'm going to take some liberties with my post on this topic. I will describe the topic of Clustering not about the industry that I work in, but an industry that I follow as a fan...professional basketball. An NBA team is made up of about 15 players but only 5 players play at a time. Because there are always 5 players on the court for each team, for a long time there has been a common description of the positions and all players have been put into one of these positions. These are:

1. Point Guard
2. Shooting Guard
3. Small Forward
4. Power Forward
5. Center

Teams are typically building their teams by trying to have their best 5 players across these 5 positions, and then having good backups at each of these positions. For some players, the position was a good match for the role that they would play, but for other players, the position did not describe their skillset and role well. Even today, some teams continue to describe players by these 5 positions and force each player into a positional role that may not be the best way to utilize the player. The game has changed and the old way of describing the players has become limiting. In order to get maximum value from each player's potential contribution to the team, some teams have recognized that a different clustering approach has started to be taken.

In recent years, some teams have taken a different approach and clustered players using different Observations. Some of the current Observations used for clustering players are as follows:

- Primary Ball Handler
- Secondary Ball Handler
- Shot Creator
- Slasher
- Athletic Finisher

- Off Screen Shooter
- Movement Shooter
- Stationary Shooter
- Versatile Big
- Post Scorer
- Stretch Big
- Roll & Cut Big

I would be really interested to see how player data could be analyzed based on these cluster groupings instead of the old positional groupings. Apparently some teams have started doing that. Evaluating the data on all players using these different cluster groupings could give new insight into the best way to value players in different roles and the best way to form a team based on a mix of players from these cluster groupings. How a team chooses to spend their money to build their team rosters could change based on this. And how a team's roster is utilized can change as well as a team may evaluate their roster and realize that a different offensive plan or defensive plan could be more effective based on the mix of players that they have from these clusters.

Professional sports gathers an incredible amount of data related to player's physical attributes (height, weight, standing reach, wingspan, vertical leap, body fat index, speed, agility, etc.) and performance (games played, minutes played, points, rebounds, assists, scoring efficiency, defensive metrics, etc.). Analyzing these large volumes of data using clusters could reveal something that gives one team an advantage over competition.

Maybe the Observations described above (the new role descriptions) are a good way to cluster the players, or maybe a different clustering approach would be more effective. In any case, I see a lot of potential for clustering to be used in sports data analytics such as what I have described here.

Reference: I used the following web-page to get information for my post: https://www.bball-index.com/offensive-archetypes/

Reply (2 likes)

Lee Lanzafame (https://classroom.emeritus.org/courses/9054/users/231975)

Apr 22, 2024

wow i didnt realise this, as someone who played basketball it would have been great to be considered something other than a forward

Reply

Ricardo Anaya (https://classroom.emeritus.org/courses/9054/users/228915)

Apr 22, 2024

great example,

but would it be usefule to have more personal info also (age, familly, financial, even criminal records) some psicological data to add to this?

any sport is not physcal, the mental state can determine future also.

Reply

Haitham Farag (https://classroom.emeritus.org/courses/9054/users/233864)

May 26, 2024

Very good example thanks Shahrod, with new insights.

Each of the observations (5 clusters) you intitally mentioned are currently broken down further by a set of skills (shot creator....etc) associated with the player role. Then depending on how each player is rated in each of those features he would be placed and ranked in each of the 5 main clusters (e.g. player X is second option shooting guard and third option point guard).

Did I get this right?

Reply

STEPHEN HUTSON (https://classroom.emeritus.org/courses/9054/users/233645)

Apr 21, 2024

In the consulting industry, where teams often comprise consultants with diverse skillsets and expertise, I can see how clustering could be really useful for the internal talent management decision making when assessing the performance of these consultants. By clustering consultants by things like skills, experience, and project history, HR teams could effectively manage and allocate resources to optimize team composition and project assignments.

Clustering could also help enable talent management to identify emerging skill gaps and guiding strategic training and development initiatives to enhance overall team capabilities. I

think it is a good data-driven approach to looking at and managing a large scale pool of employees which is difficult to manage individually after a certain scale is reached

Reply

(http

Lee Lanzafame (https://classroom.emeritus.org/courses/9054/users/231975)

Apr 22, 2024

Career development (skill gaps) like you mentioned is a great use case

Reply

(https:/

Lee Lanzafame (https://classroom.emeritus.org/courses/9054/users/231975)

Apr 22, 2024

I work at a telecommunications company. We often try to predict if a customer is likely to churn. We have centralised customer data that gets fed into a k-means clustering algorithm that results in a likelihood of a customer churning.

Some of the data that contributes to this is a customers age, demographics, location (regional / metro), number of recent call centre interactions, billing info, avg download and upload speed and also number of dropouts.

My company currently has a large datalake where we leverage machine learning in many different teams; network optimization, fraud detection, customer segmentation are some of the areas that benefit from this.

Reply

(http

Ricardo Anaya (https://classroom.emeritus.org/courses/9054/users/228915)

Apr 22, 2024

churn is too subjective,

but time variables can help (time being a customer) time between calls, time between complains

Time of service is up, time to adress an issue.

Reply

Chris Cosmas (*He/Him*) (https://classroom.emeritus.org/courses/9054/users/226607)

Apr 22, 2024

The main objective of a Central Bank is to assess the monetary framework of the economy and assess the robustness of the financial system to withstand shocks in markets regionally or internationally. This is done mainly by monitoring Banks, Insurance Companies and other licensed financial institutions. There are specific departments which monitor the financial stability of the local economy and the resilience of banks.

In this scenario the observations are individual banks with each hundred of data points that are reported to the Central Bank. Such as the ECB and local European banks which provide the ECB with solvency and liquidity metrics defined by legislation such as Basel II.

Clusters can be created by looking at the exposure of banks to certain industries such as the real estate industry as a percentage of their total assets this can be evaluated in relation to total liquid assets such as cash, market securities, short term accounts receivable, etc...

This can give regulators an understanding of the different investment profiles and strategies employed by banks. This allows the regulator to monitor the risk profile of banks and formulate fitting regulation which is aligned with the invested industries potential risks making sure banks conduct themselves in a correct manner.

Reply (1 like)

Jignesh Dalal (https://classroom.emeritus.org/courses/9054/users/229173)

Apr 23, 2024

That's a very insightful observation about the use of clustering within the context of a Central Bank. It seems particularly valuable in understanding the exposure and risk profiles of various banks based on their investment strategies. This approach could indeed enhance the precision of regulatory frameworks by tailoring them to specific risk profiles, which could be instrumental in preventing systemic risks.

Given this potential, do you think there are also possibilities to use clustering to monitor the behavior of these institutions over time? For instance, could clustering help in identifying trends or shifts in investment strategies that might indicate a movement towards higher risk portfolios, especially in volatile market conditions?

Also, considering the critical nature of real-time data in financial stability, how effective do you think current data collection methods are in providing the necessary granularity and timeliness to support this kind of analytical approach?

Edited by Jignesh Dalal (https://classroom.emeritus.org/courses/9054/users/229173) on Apr 23 at 1:34am

Reply

Jignesh Dalal (https://classroom.emeritus.org/courses/9054/users/229173)

Apr 22, 2024

Q) consider the types of clustering that can be useful to your current industry. Can you identify specific areas in your organization that would benefit from clustering?

Currently, I work in telecommunication industry. Where focus in on selling Internet, Telephone, TV, Mobile Wireless services and add on features like Disney, Netflix etc.

Below, I'll explain the concept of clustering, identify key areas in the telecommunication industry that would benefit from this techniques, and provide recommendation on its application.

Clustering is a type of unsupervised learning method used to group a set of objects in such a way that objects in the same group are more similar to each other that those in other group.

Application of clustering in telecommunication for a marketing campaign.

Clustering techniques will allow marketing team to identify distinct groups within its customer base and tailor marketing strategies accordingly. Here are few marketing-focused applications of clustering in telecommunication.

Application Area	Clustering Method	Purpose of Clustering	Benefits	Hypothetical Benefits
Customer Segmentation for Personalized Marketing	K-means, Hierarchical	To segment customer based on behaviour, demographics and usage patterns.	Enables personalized marketing, improves engagement, and increases sales conversation.	Identification of five distinct customer segments leading to a 30% increase in targeted campaign response rates.
Campaign Optimization	K-mean	To analyze the effective of different	Optimize marketing efforts, focuses on	Redirection of budget resulting in a 25% cost

Application Area	Clustering Methods	Purpose of Clustering	Benefits	Hypothetical Benefits
	s	marketing campaign across customer cluster.	effective tactics and channels.	reduction while maintaining campaign effectiveness.
Customer Lifetime Value Prediction	Hierarchical	To group customer based on their lifetime value.	Prioritized high value customers, customizes loyalty programs.	Development of tier based loyalty programs increasing customer retention by 20%in the high value segment.
Geographic Targeting	K-means	To segment markets based on geographic data and regional preferences.	Enhances regional marketing strategies, optimizes resource allocation.	Customized regional campaigns boosting sales by 15% in underperforming areas.

Reply (1 like)

Chris Cosmas (*He/Him*) (https://classroom.emeritus.org/courses/9054/users/226607)

Apr 24, 2024

Thank you for your input Jignesh it is interesting to see different scenarios in which clustering might be applied as well as the perceived benefits of using statistical techniques in a business application.

Is there a reasoning behind the choice of the clustering method for each scenario?

Reply

Ricardo Anaya (https://classroom.emeritus.org/courses/9054/users/228915)

Apr 22, 2024

I work in the Cellular industry

where I need to manage Customers that include

Wirress Operators (data of interest: Spectrum holding band, and amount of spectrum) that varies per country and many countries have multiple operators.

OEM that haveSpecific set of bands per product, accorfing to market

Internal Products that have specific bands

bands include Band number, bandwith and technology

technology can be 2G, 3G, 4G, 4.5G and 5G

4.5G and 5G include combinartion from 2 bands up to 7.

we can cluster bands per range:

Low band mid band and High Band

Then we can clusert them by Product Support

We can clsuter them by REgion (that is happening)

Reply (1 like)

(http

Ricardo Anaya (https://classroom.emeritus.org/courses/9054/users/228915)

Apr 22, 2024

I want towork in this example to cluster them in new ways to predict volume, to manage prices by demand,

to group products into set of bands to offer them as a bundle, I wan to analize the data to have backup to present to management.

Reply

(http

Manjari Vellanki (https://classroom.emeritus.org/courses/9054/users/231480)

Apr 25, 2024

Hi Ricardo-

Interesting example! In this scenario, what are the key variables that impacts the clustering process?

Reply

Mariana Flores (https://classroom.emeritus.org/courses/9054/users/237198)

Apr 23, 2024

Cluster analyses have wide range of application across many industries due to its multifaceted ability to discover insights from complex sets of data. These techniques are useful when partnering with small business owners for pattern recognition and data summarization projects. K-Means Clustering with Principal Component Analysis (PCA) is one example which is very useful for personalization. When identifying customer segments and grouping sub-segments of customers based on purchase behavior, preferences, or demographics to target for personalized offers it provides a more effective manner for upsell and cross-sell initiatives of various products and/or services. Natural Language Processing (NLP) is another clustering technique which can be utilized for text summarization of product recommendations and to create an NPS score through sentiment analysis. The adaptability and extensive range in application of cluster analyses makes these techniques a crucial tool across many industries by helping to intelligently guide business decision-making and inform both strategic choices and resource allocation.

Edited by **Mariana Flores (https://classroom.emeritus.org/courses/9054/users/237198)** on Apr 23 at 3:59am

Reply

Priscilla Annor-Gyamfi (https://classroom.emeritus.org/courses/9054/users/226376)

Apr 23, 2024

In my current industry, the arts and crafts, there are specific areas in my organization that would benefit from clustering. These areas are.

1. Customer Clustering: Customer Grouping: Clustering aids in recognizing different customer segments according to their shopping habits, demographics like age, gender, and location, or other relevant characteristics. This approach helps in making well-informed choices regarding marketing strategies, promotional activities, product performance, and customer service approaches to cater to the unique requirements of each segment.

2. Inventory Management: Inventory Organization: Clustering assists in refining our inventory management system by categorizing products based on their demand, seasonal variations, and sales performance. By grouping products with similar demand patterns, we improve our capacity to predict demand accurately, maintain appropriate stock levels, and decrease the chances of shortages or surplus stock. This method allows us to manage resources efficiently and build trust and confidence with our customers.
3. Effectiveness in Operational Processes: Clustering can help optimize various operational processes within our Industry. For instance, grouping employees based on their skill sets, experience and performance metrics can help identify opportunities for growth in terms of offering promotions and incentives(bonuses) as well as knowledge sharing.
4. Spotting Trends: Clustering can help spot ongoing and upcoming trends in the arts and crafts sector, particularly in beaded accessories. By grouping items by features like color, theme, or design elements (like materials), we can pinpoint trends growing in favor with customers. This allows us to adapt our product creation and marketing strategies to make the most of these trends.
5. Personalized Recommendations: Through clustering, we can study customers' past purchases and trends to create customized product suggestions that match each customer's tastes. By grouping similar purchase histories, the art industry can recommend products likely to appeal to individual customers, enhancing their shopping experience and satisfaction."

Reply

Ahmad Abu Baker (https://classroom.emeritus.org/courses/9054/users/234460)

Apr 24, 2024

Hello Priscilla Annor-Gyamfi,

It’s great to see your strategic use of clustering in the arts and crafts industry. Your approach to customer segmentation is spot-on for enhancing marketing strategies and customer satisfaction. Similarly, optimizing inventory management through clustering will help maintain the right stock levels and avoid over or understocking.

Using clustering to group employees can improve operational efficiency and foster a growth-oriented workplace culture. Additionally, spotting trends and offering personalized recommendations are innovative ways to stay ahead in the market and enhance the customer shopping experience.

I'm excited to see the positive impacts these initiatives will have on your operations and customer relations.

Best regards,
Ahmad Baker

Reply

Haitham Farag (https://classroom.emeritus.org/courses/9054/users/233864)

May 26, 2024

Spotting Trends is an insightful use case, a valuable dimension. Thanks, Priscilla for highlighting.

Reply

Todd Engle (https://classroom.emeritus.org/courses/9054/users/228910)

Apr 23, 2024

I'm a Project Manager by profession and I've always disliked the very linear way Project Management Offices calculate project risk. Risk on a project is usually set by looking at specific variables, such as effort, technology, and budget with very little thought to historical data.

I could see there could be a great improvement in the model by assessing Project Risk by:

- Clustering projects based on factors like budget size, team size, project complexity, **and** historical performance data.
- Analyze the clusters to identify common characteristics of projects with high or low risk.
- Use these insights to develop targeted risk mitigation strategies for different project types.

It would also be a great tool, especially for large projects to help with resource allocation and a well-developed resource hiring plan. In my current position it was clear that a resource plan was not developed and it has caused massive delays, high turn-over and a large loss of tacit knowledge. I think this could be accomplished by…

- Cluster projects based on required skillsets, software needs, and timelines.
- Analyze the clusters to understand the resource demands for each project type.
- Use this information to allocate resources efficiently and avoid bottlenecks.

In my 30 year experience, these types of analysis are very important, but rarely done during the feasibility stage of the project. This type of analysis, if properly actioned, could reduce a lot of the project factors that cause them to have costly budget and timeline overruns.

Reply

(http

Chris Cosmas (*He/Him*) (https://classroom.emeritus.org/courses/9054/users/226607)

Apr 24, 2024

Hello Todd,

It's a very good idea, organisations are now understanding the value of data driven decision making. This approach enables decision makers to identify patterns in projects giving insights on the required resources to successfully complete projects avoid the impact of resource constraints and be well prepared due to the indications given by past projects. It is also interesting to see how clustering can be fit to different tasks and objectives focusing on the characteristics which are most relevant on the user's perspective.

Reply

(http

Swati Sharma (https://classroom.emeritus.org/courses/9054/users/236938)

Apr 30, 2024

Hello Todd: Your suggestion to use clustering analysis for project risk assessment and resource allocation in project management is innovative and addresses common challenges. By clustering projects based on factors like budget size, team size, and historical performance data, you propose a more holistic approach to risk assessment. This allows for the identification of common characteristics among projects with high or low risk, enabling targeted risk mitigation strategies. Additionally, your recommendation to cluster projects based on required skill sets and timelines for efficient resource allocation aligns with best practices. By analyzing clusters to understand resource demands, organizations can optimize resource utilization and mitigate costly budget and timeline overruns.

Reply

(https:/

Ahmad Abu Baker (https://classroom.emeritus.org/courses/9054/users/234460)

Apr 24, 2024

Hey team,

After diving into the latest module on clustering, I've been thinking a lot about how we can apply these concepts to shake things up in my current industry of business consulting, specifically transformation delivery. Here are a few ideas I think could really benefit our day-to-day operations and client strategies:

1. Tailoring Client Solutions Through Segmentation: Imagine if we could group our clients into clusters based on factors like their industry, company size, or specific challenges they face. We could use these insights to customize our advice and resources more sharply, making sure our strategies hit the mark for each type of client.

2. Building Stronger Project Teams: We all know that the right team can make or break a project. What if we used clustering to help us put together teams? By grouping consultants who mesh well or share similar skills, we could ensure smoother collaboration and even spark some new, innovative ideas for our projects.

3. Proactive Risk Management: Nobody likes surprises, especially the risky kind. By using past project data, clustering could help us spot potential pitfalls before they happen, letting us prepare better and tackle risks head-on with strategies that we know work for those particular scenarios.

4. Spotting Red Flags in Projects: Sometimes small things can signal big problems. Clustering can be our early warning system, helping us notice when a project isn't tracking as it should. This way, we can jump in and fix issues before they grow into bigger challenges.

5. Smarter Resource Allocation: We're always trying to do more with less, right? Clustering can guide us in mapping out our project demands and matching them with available consultants. This means better planning, less scrambling, and more efficient use of everyone's time.

I’m really excited about the potential here to not just improve our efficiency, but also enhance the bespoke strategies we pride ourselves on delivering to our clients. Keen to hear your thoughts and maybe kick off some projects where we can test these ideas out!

Looking forward to some lively discussions!

Best,
Ahmad Baker

Edited by **Ahmad Abu Baker (https://classroom.emeritus.org/courses/9054/users/234460)** on Apr 24 at 1:05pm

Reply

(http

Mhelissa Yayalar (https://classroom.emeritus.org/courses/9054/users/233590)

Apr 28, 2024

Hi Ahmad,

I find your application of clustering in project management area very interesting. In my current role as a program manager, managing many projects at once is very challenging. As you noted regarding proactive risk management, planning for risk or trying to anticipate risks is an important part in making the project is successfully implemented. I read a blog by Paul Boudreau about this related to this topic in project management. Paul specifically touched on using unsupervised learning to cluster risks for developing risk mitigation strategy and also finding a risk that was overlooked. For instance, using clustering algorithm to classify the number of story points for each story can help project managers discover if there's a risk in properly completing the project because perhaps points did not accurate align with the level of effort that required for the story to be considered complete.

Reference:

Paul Boudreau. “How Unsupervised Learning Algorithms Help Project Management.” AI IQ Blog, October 09, 2023. **https://www.projectmanagement.com/blog-post/75205/how-unsupervised-learning-algorithms-help-project-management#_=_ (https://www.projectmanagement.com/blog-post/75205/how-unsupervised-learning-algorithms-help-project-management#_=_)**, 27 April, 2024.

Reply

(https:/

Swati Sharma (https://classroom.emeritus.org/courses/9054/users/236938)

Apr 26, 2024

Hello all: In my past i was working in a e-commerce company, i was the data analyst for my team to provide insights about how the revenue is coming. the leader really struggled to identify how to retain customers and make marketing strategy to make customers buy more. Today after learning the concept of clustering i would used different clustering methods Eg K-Means clustering to identify frequent buyers, occasional shoppers, high spenders, etc and also group products based on similarities in their attributes, such as category, price, brand, and features.This will help the leaders to make making strategies suitable for different customers based there purchase history.

Reply

Mhelissa Yayalar (https://classroom.emeritus.org/courses/9054/users/233590)

Apr 27, 2024

In my job industry, which is in the aerospace industry, using clustering algorithm drives actionable results that yields to revenue growth. My company's structure provides several areas where clustering can enhance collaboration, innovation, and efficiency. Whether it's within specific business units or across the entire organization, clustering plays a vital role in shaping the future of aerospace industry.

The specific business unit I support leverage clustering to optimize supply chain logistics, improve global operations, and enhance customer support. When it comes to optimizing sales, supply chain logistics, improving global operations, and enhancing customer support, I think the following clustering can be beneficial:

1. K-Means Clustering is used for Sales & Marketing for customer segmentation. For instance, we can group customer similarities by products, regions, and portfolio size clusters. The use of K-Means enables us to tailor sales lead strategies for different customer segments, optimize inventory management, and allocate services resources.
2. Hierarchical Clustering helps in global operations optimization by organizing airline data into a tree-like structure of clusters, allowing as to explore different levels of customer segmentation. For instance, we can conduct deeper dive on most common parts are utilized by each region. This method provides decisions related parts services allocation across our global operations.

Reply

Timothy Andrew Ramkissoon (https://classroom.emeritus.org/courses/9054/users/226697)

May 1, 2024

Hi Mhelissa,

Something I would consider is, to ensure that the features/variables used for segmentation are relevant and meaningful. Consider collaborating with domain experts to identify critical attributes that impact revenue growth, supply chain efficiency, and customer satisfaction. For example, in sales and marketing, besides product and region, explore additional features such as customer behaviour, purchase frequency, or lifetime value.

Reply

Koffi Henri Charles Koffi (https://classroom.emeritus.org/courses/9054/users/208039)

Apr 29, 2024

In an organization to detect and assign right to employee to login or accessing resource of the company
base on the current status and the position of employee in employee in the organization

- **The feature or attribute of employee**
  position of the employee(junior, senior , director, president)
  status of the employees (new join, current employee, moving to new department, leaving the company)

here we can use the K-mean algorithm with initial k = 16

Here we can have around 16 clusters, with each cluster the combination of each feature .

- For each cluster assign a score . with generated score we can give access Right to resource available in the company
- The score for each cluster can be used to detect malicious user login activity

Edited by Koffi Henri Charles Koffi (https://classroom.emeritus.org/courses/9054/users/208039) on Apr 29 at 1:19am

Reply

Timothy Andrew Ramkissoon (https://classroom.emeritus.org/courses/9054/users/226697)

May 1, 2024

As someone who works in the Asset Integrity Management space, there are some considerations for how clustering may benefit my organisation.

This can be useful for asset life extension studies. Comparing clusters of aging assets can reveal trends related to their remaining useful life. BY analyzing maintenance records, failure rates, and degradation rates, we can identify clusters of assets that require proactive

interventions. This can be used to prioritize maintenance efforts, extend the life of critical components, and optimize asset utilization.

Comparing clusters of subsea equipment based on their conditions can assist in developing condition-based maintenance programs. We can allocate resources more carefully by focusing on clusters with higher risk or imminent maintenance needs. This approach minimizes downtime and reduces overall maintenance costs.

Also, cluster analysis can help identify patterns in production data from subsea fields. By grouping similar assets based on performance metrics (such as production rates, downtime, or efficiency), we can uncover insights. Comparing clusters allows us to understand which operational practices lead to optimal production levels.

Reply

(https:/

Gustavo Santana (https://classroom.emeritus.org/courses/9054/users/120927)

May 1, 2024

Working in a company that offers English Classes for companies by mobile app, including private classes with English professors, it would be interesting to cluster the student's data and understand the different behaviors or student personas we attend.

One interesting aspect is that observing students from different regions makes it possible to design some patterns, for example, Asian students tend to be more disciplined and study for more hours than their Western peers, Latin Americans tend to study sporadically sometimes in the morning or later at night.

Reply

(http

Isabella Tockman (https://classroom.emeritus.org/courses/9054/users/207395)

May 3, 2024

I came to the US to learn English, and my first roommate happened to be a Japanese girl. We shared the room, and every morning when I woke up, she was already deeply focused on her studies. Your point is very true!

Reply (1 like)

Isabella Tockman (https://classroom.emeritus.org/courses/9054/users/207395)

May 3, 2024

I work in the construction industry, and one of the costs that hurt us the most is the cost of insurance since accidents are very common, and we face many lawsuits. So, thinking about how we could apply clustering techniques across various dimensions of our organization to effectively identify and mitigate risks associated with work accidents in the construction field is a great idea. We could cluster construction sites based on historical accident data, environmental conditions, and project characteristics to pinpoint future accident hotspots. Additionally, we could cluster employees based on their job roles, experience levels, and training records to assess task-specific risks and tailor safety interventions accordingly. This could help us assess the safety culture within work teams, fostering a positive safety culture with targeted interventions such as safety leadership training and team-building exercises. Furthermore, we could integrate real-time monitoring technologies for proactive intervention to prevent future accidents as they occur by detecting anomalies and triggering immediate safety warnings or emergency response protocols. This comprehensive future approach, leveraging clustering across our organization, could ensure a proactive and targeted strategy to reduce employee work accidents and enhance workplace safety in the construction field.

Reply