**“Customer insights and strategy optimizations”**

Petr Datsenko

Department of Art and Science,

Middle East Technical University

STAT112 Introduction to Data Processing and Visualization

**Prof. Ceylan Talu Yozgatlıgil**

**December 8, 2023,**

**Introduction**

In the vast mosaic of American states, where each region adds a unique hue to the canvas of consumer behavior, our project, "Customer Insights and Strategy Optimizations," sets out on a journey of discovery. Imagine this endeavor as an artistic exploration, where we delicately brush through the nuances of customer demographics and purchasing behaviors, seeking patterns and rhythms that echo across the diverse landscapes of the United States.

Our exploration begins with understanding how different age groups and genders, akin to the varied colors on a painter's palette, correlate with specific purchasing patterns. The map is visualized in “canvas,” where each state shows the difference between purchasing categories. It's an attempt to decipher the artistic expression of consumer choices within the picture of demographics.

Transitioning to the changing size of consumer behavior, our project employs a bubble chart as our artistic tool. This bubble chart unveils the intricate patterns of purchase amounts, reflecting the vibrant sizes, colors, and ages that sway customers' buying decisions across the diverse landscapes of the United States that can be seen in the dashboard.

Delving deeper, we investigate the influence of discounts and promotions like market vendors enticing passersby. Picture a bustling marketplace where different demographics respond uniquely to these promotions. It's a vibrant dance where values play the tune, and customers move to the rhythm of their preferences.

Venturing into loyalty and subscriptions, our inquiry mirrors a journey where customers subscribe to the loyalty narrative. Imagine loyalty as a cherished melody, echoing through the frequencies of purchase frequency, review ratings, and past purchases, telling stories of satisfaction and commitment.

In the logistical journey of choices, the influence of shipping types and payment methods mirrors a review rating. Different varying modes of delivery and payment favor review ratings. It creates a topography of consumer choices that affect the review ratings and repeat purchases, similar to the preferred routes in a traveler's journey.

Our exploration concludes by unraveling patterns in purchase amounts, a grand finale where each state contributes to the overarching melody of consumer value. The canvas is rich with insights, not just in the frequency of purchases but in the enduring value that each customer brings to the marketplace.

As we embark on this artistic exploration of data, our brushes are the algorithms, and our canvas is the diverse landscape of American consumer behavior. The strokes and patterns we uncover promise insights and the potential to optimize strategies across this grand canvas. Welcome to the artistic expedition of "Customer Insights and Strategy Optimizations" across the enchanting states of the USA.

**Data Processing**

In the meticulous choreography of data preparation, our tool for tidying up was the familiar Excel spreadsheet. It's like wielding a digital broom to sweep away any cobwebs that might obscure the clarity of our insights. Allow me to share a glimpse of this tidying endeavor.

In the file “country.xlsx,” our eagle-eyed observation detected duplicity within the list of states. Like a meticulous curator refining an art collection, the duplicated conditions were identified and promptly removed. Imagine this process as a curator spotting identical paintings in an exhibition, ensuring that each piece stands distinct and contributes uniquely to the overall narrative.

So, as we embark on the journey of data cleaning, envision Excel as our trusty toolkit and remove duplicated states as a gentle pruning, ensuring that our dataset blossoms into a clean and uncluttered canvas ready to absorb the vivid strokes of analysis. Welcome to the initial steps of our data journey, where clarity begins with the gentle sweep of a digital broom.

**Exploratory Data Analysis**

To conduct exploratory data analysis (EDA), let's start by providing descriptive statistics and interpretations for crucial variables in the dataset. We will focus on gaining insights into the critical variables' central tendencies, distributions, and patterns.

**1. Variables:**

* **Age:**

Descriptive Statistics: Mean, median, standard deviation, minimum, maximum.

Interpretation: Assessing the central tendency and spread of customer ages, providing insights into the age distribution.

* **Gender:**

Descriptive Statistics: Count and percentage of each gender.

Interpretation: Analyzing the gender composition to understand the demographic representation.

* **Purchase Amount (USD):**

Descriptive Statistics: Mean, median, standard deviation, minimum, maximum.

Interpretation: Examining the central tendencies and spread of purchase amounts to identify common spending patterns.

* **Review Rating:**

Descriptive Statistics: Mean, median, standard deviation, minimum, maximum.

Interpretation: Assessing the average satisfaction level of customers based on review ratings.

* **Subscription Status:**

Descriptive Statistics: Count and percentage of subscribers.

Interpretation: Understanding the prevalence of subscription services and potential implications for recurring revenue.

* **Frequency of Purchases:**

Descriptive Statistics: Mean, median, standard deviation, minimum, maximum.

Interpretation: Evaluating the average frequency of customer purchases, providing insights into customer engagement.

These descriptive statistics offer an initial understanding of the dataset's critical variables. Further analysis, such as visualizations and statistical tests, can be conducted to explore relationships and patterns within the data. Remember that these interpretations serve as a starting point, and a more in-depth EDA would involve a comprehensive exploration of each variable and its interconnections.

**2. Questions for analyzing data:**

* **Customer Demographics and Purchasing Patterns:**

How do gender and location correlate with specific purchasing patterns, such as purchase amounts and preferred categories?

A map of the united states

Description automatically generated

In addressing the inquiry regarding the correlation between gender, location, and purchasing patterns, a meticulous approach was taken to craft an immersive and insightful visual experience. The result was the creation of an interactive map that serves as a portal into the intricate world of consumer behavior.

This carefully designed interactive map invites users to embark on a journey of exploration. As you delicately hover over each state, a dynamic and illuminating stacked bar chart materializes, offering a layered visual narrative that transcends traditional data analysis. This interactive masterpiece provides a visual feast and a powerful analytical tool, unraveling the nuanced dance of preferred item categories based on gender.

The visual narrative unfolds with each hovering motion, offering a symphony of colors and shapes representing the heartbeat of consumer choices. It is akin to peeling back the layers of a captivating story, where each state unveils its unique chapter in the grand tale of purchasing patterns.

The beauty of this visual exploration lies in its ability to capture the essence of consumer choices across diverse locations. It goes beyond the superficial, delving into the intricacies of gender-specific preferences and their manifestation in distinct categories. This dynamic representation paints a vivid picture of the kaleidoscopic landscape of purchasing patterns, providing viewers with a nuanced understanding of the interplay between gender, location, and specific purchasing behaviors.

Furthermore, this visual journey extends beyond mere aesthetics; it is a treasure trove of insights into purchase amounts and the categories that hold sway in each state. The varying hues and heights of the stacked bars tell a story of consumer preferences that transcends conventional narratives. It's not just about what is purchased; it's about the rhythm and melody of choices that resonate uniquely in different corners of the map.

Through this innovative and meticulously crafted visual exploration, a deeper understanding emerges—a narrative that goes beyond the numbers and charts. It is a testament to visualization's power in unraveling the complexities of consumer behavior. This interactive map and stacked bar chart duo are a testament to the synergy between design and data, providing a nuanced perspective on the interwoven threads of gender, location, and purchasing patterns. Welcome to the immersive world where data comes to life, offering insights beyond the surface.

* **Age Trends and Product Preferences:**

How do factors like the color, size of purchased items, and age vary with customer demographics?

A screenshot of a computer screen

Description automatically generated

In delving into the exploration of age trends and product preferences, a deliberate and insightful approach was employed, leveraging the sophisticated capabilities of a meticulously crafted packed bubble chart. This visual masterpiece serves as a gateway to unraveling the intricate nuances of customer demographics and their dynamic interactions with product preferences.

Utilizing a packed bubble chart was not merely a choice but a strategic decision to elevate the storytelling aspect of data. This visually arresting representation goes beyond the conventional, offering a vivid and nuanced response to the inquiry. As we embark on this visual journey, the chart unfurls like a canvas of possibilities, each bubble encapsulating a story waiting to be deciphered.

The narrative begins by exploring the top 5 colors, a palette that mirrors consumers' tastes and preferences. The chart's dynamic composition vividly illustrates the dominance of shades of gray and white, emerging as the frontrunners in the symphony of color preferences. The visual feast extends into product sizes, mainly throw sizes, where choices take on a fascinating vibrancy.

The size of each bubble, strategically placed within the chart, serves as more than a visual cue; it becomes a storyteller in its own right. Each bead is a microcosm of customer age trends, creating a dynamic interplay between the demographic dimensions of age, color, and size preferences. The chart becomes a visual orchestra, where more minor bubbles dance to the tune of younger generations, and larger ones resonate with the likes of older age groups.

This visualization is not just a snapshot of preferences; it is a visual narrative that unfolds layers of insights. The dominance of specific colors and sizes tells a story of what customers buy and how these choices align with their age demographics. It explores the interwoven threads of consumer behavior that extend beyond the surface, revealing patterns that might elude traditional analyses.

As we immerse ourselves in this visual tapestry, the packed bubble chart becomes a portal into the intricate dance between age, color, and size preferences. It's not merely a chart; it's a visual journey that invites us to explore the subtleties and complexities of consumer choices within the dynamic landscape of age demographics. Welcome to a realm where data transcends numbers and transforms into a vibrant tapestry of insights.

* **Impact of Discounts and Promotions on Sales**

How influential are discounts and promotional codes in influencing customer purchasing decisions? Do specific demographics or customer segments respond more to these promotions?

A graph of blue and orange bars

Description automatically generated

Embarking on the enlightening journey to unravel the profound impact of discounts and promotions on the intricate tapestry of customer purchasing decisions, I meticulously navigated the complex labyrinth of consumer behavior using the immersive visual tool of a side-by-side bar chart. This deliberate choice aimed to present information and craft an engaging narrative that would encapsulate the nuanced dynamics of promotional strategies, stretching far beyond the surface level of demographic and segment considerations.

As the canvas of our exploration unfolded, the side-by-side bar chart became more than just a graphical representation; it metamorphosed into a visual symphony, orchestrating the intricate dance of categories responding to the ebb and flow of promotional deals. The beauty lies in the static presentation and the dynamic revelations that a cumulative sum of previously purchased amounts can unveil—a temporal dimension that breathes life into our analysis.

The strategic choice to delve into the cumulative sum over time becomes a masterstroke, allowing us to trace the enduring influence of values as they echo through the corridors of consumer choices. The visual narrative takes on a rhythmic quality, akin to a melody composed of data points that harmonize to reveal the enduring impact of promotions on the consumer psyche.

As we traverse the visual landscape, an intriguing story unfolds—one of clear dominance and unexpected preferences. The data paints a vivid picture of category preferences, with clothing emerging as a towering presence and a beacon in the consumer choice hierarchy. Juxtaposed against this dominance is the revelation that outerwear, intriguingly, appears as the least favored category—a subtle nuance that adds layers to our understanding of consumer preferences.

However, the data doesn't merely rest in the realms of expected outcomes; it invites us to explore the unexpected. A heightened affinity for purchasing more when the siren call of discounts is momentarily absent introduces a captivating twist to our narrative. It's a paradoxical dance where consumers, without explicit incentives, reveal a penchant for increased engagement. This revelation challenges preconceived notions and opens doors to a more profound exploration of the intricacies of the consumer decision-making process.

This comprehensive analysis transcends the conventional boundaries of demographics and segments. It's not just about identifying which categories rise to dominance; it's about deciphering the intricate dance between promotions, consumer choices, and the ever-evolving landscape of preferences. Our side-by-side bar chart transforms into a dynamic visual saga, where each element is a brushstroke contributing to the vibrant canvas of insights. Welcome to an immersive exploration where data transcends the mundane, transforming into a captivating narrative of consumer behavior that beckons us to delve deeper into the layers of promotional dynamics.

* **Influence of Shipping Type and Payment Method on Customer Choice**:

How do the shipping type and payment method preferences vary among customer segments? Does the choice of shipping type affect the review ratings or repeat purchase behavior?

A screenshot of a graph

Description automatically generated

In the comprehensive exploration of shipping type and payment method preferences, the heatmap emerges as the quintessential guide, unraveling the intricate tapestry of customer choices. Within this dynamic representation, every cell serves as a unique testament, bearing witness to the nuanced relationship between review ratings, shipping types, and payment methods. The subtle dance of colors within the heatmap unveils a mosaic of customer sentiments—lighter shades symbolizing lower ratings, while deeper hues signify higher satisfaction.

As we delve further into this correlation across diverse customer segments, a meticulous dissection occurs, emphasizing a gender-based analysis. Noteworthy insights surface, revealing an intriguing pattern: male customers tend to assign lower scores than their female counterparts. This tailored exploration transcends the general landscape, offering a granular understanding of how shipping preferences, payment methods, and review ratings intertwine.

The narrative within this detailed analysis extends beyond a mere visual representation; it becomes a nuanced tale of preferences, perceptions, and patterns. Through this tailored approach, we spotlight the variations across diverse customer segments, enriching our comprehension of the multifaceted dynamics shaping the customer experience.

* **Customer Feedback and Business Strategy:**

How do review ratings correlate with repeat purchases, subscription status, and demographic factors? Can customer feedback be used to predict future purchasing behavior or preferences?

A network of dots and lines

Description automatically generated

Embarking on exploring intricate interconnections among review ratings, repeat purchases, subscription status, and demographic factors, my chosen visualization tool, the network diagram, stands as a beacon of clarity. This selection arises from its unique capability to vividly illustrate correlations within various categories, where the size of each node becomes a tangible representation of significance.

As we navigate the network diagram's visual narrative, the subscription status divergence across diverse locations unfolds like a captivating story. Each location node's size serves not merely as a visual cue but as a profound symbol intricately tied to the average rating. The resulting graphical representation assumes a form akin to a symmetric star, a visible emblem of the consistent relationship between subscription status and average rating across different regions.

This nuanced exploration transcends the surface, peeling back layers of customer feedback, subscription status, and review ratings. Insights from this detailed analysis transcend immediate observations, offering a deeper understanding of their potential implications for future purchasing behavior and preferences. The network diagram becomes more than a visual aid; it transforms into a storyteller, unraveling the complexities that shape customer interactions and laying the groundwork for strategic decision-making.

* **Analysis of Purchase Amounts and Customer Value:**

What are the patterns in purchase amounts across different customer segments, product categories, and locations? How do these patterns relate to the frequency of purchases and customer lifetime value?

A graph with blue squares

Description automatically generated

In navigating the analysis of item purchases, my tool of choice, the waterfall chart, becomes a lens through which we decode the prevalence of assets and unravel insights into the nuanced distribution of purchase amounts. These charts, resembling a mosaic of distinct bars representing items, not only shed light on their relative differences but also seamlessly provide a cumulative total of purchased items. In this analytical journey, I intentionally opted for the most fitting metric—purchase amount—to ensure a comprehensive exploration.

As we delve into the intricacies presented by the chart, a discernible pattern emerges: the ascending trajectory of the diagram unfolds like a narrative, suggesting a nearly uniform distribution of purchases across all items. This pivotal observation signifies a notable absence of significant disparities in purchase quantities. It affirms a captivating reality: each item experiences nearly equal customer popularity.

This meticulous analysis transcends the immediate observations, addressing the broader question of understanding patterns in purchase amounts. It offers a valuable vantage point across diverse customer segments, product categories, and locations. These insights become instrumental in unraveling the intricate relationship between these patterns and influential factors, such as the frequency of purchases and customer lifetime value. In essence, this analysis becomes a compass, guiding us through the labyrinth of consumer behavior and providing a compass to navigate strategic decisions.

* **Relationship between category and type of shipping:**

How do categories relate to shipping types and purchases? Is there a connection between choosing a specific category and a particular type of delivery?

A blue and orange lines

Description automatically generated

I sought clarity through the insightful Sankey diagram, delving into the labyrinth of connections among item categories, shipping preferences, and purchasing behavior. This visual representation masterfully positions classes on the left and delivery types on the right, seamlessly connected by lines whose width varies, reflecting the corresponding purchase amounts. The consistent number of connections between each category and delivery type paints a vivid picture of a balanced distribution of purchased items, forming the backbone of our consumer landscape.

I deliberately highlighted three distinctive orange lines in my quest for a more detailed exploration. These lines act as beacons, spotlighting higher-value transactions within the intricate web of connections. This strategic emphasis on specific pathways allows for an in-depth analysis, revealing nuanced insights into the correlations between distinct item categories and their preferred delivery methods. The result is a rich tapestry of understanding, providing a comprehensive view of the subtle dynamics that shape item selection and the preferences governing the delivery journey.

**Dashboard**

I crafted a comprehensive dashboard to gain a holistic perspective on the intricate interplay of colors, sizes, ages, locations, shipping types, payment methods, and purchase amounts. This dashboard encapsulates critical insights from "Customer Demographics and Purchasing Patterns," "Age Trends and Product Preferences," "Influence of Shipping Type and Payment Method on Customer Choice," and "Impact of Discounts and Promotions on Sales."

Through seamless interactions within the dashboard, we unveil unique trends across diverse states, providing a panoramic view of how these variables harmonize to shape customer behaviors and preferences.

A screenshot of a computer screen

Description automatically generated

Embarking on an exploration through this dashboard opens up a realm of diverse trends waiting to be uncovered. A simple click on a state in the map unravels a wealth of insights, allowing us to delve into:

A screenshot of a computer

Description automatically generated

In our expansive dashboard, noteworthy distinctions emerge, presenting an impressive ability to efficiently navigate and comprehend extensive data specific to various states. This feature underscores the effectiveness of our comprehensive reporting approach.

**Conclusion and discussion**

We've uncovered hidden treasures in a vast sea of data as we wrap up our journey through the "Customer Insights and Strategy Optimizations" project. We have dived into the numbers using various tools like stacked bar charts, packed bubble charts, and more to understand how customers tick, what they prefer, and how we can improve their experience.

Looking at who our customers are and what they buy, we found an excellent mix across different age groups and genders. The stacked bar charts gave us a visual feast, showing the interesting patterns in what folks like to buy.

Regarding seasons and what products shine during different times of the year, we traded the usual suspects for bubble charts. These bubbly visuals gave us a playful way to see how much love each product category gets during different seasons. It's like watching the seasons unfold through what people love to purchase.

Discounts were a big player in the customer shopping game, and we laid it all out in side-by-side bar charts. These charts spilled the beans on how discounts can sway customer decisions and make shopping even more exciting.

For the loyal bunch who signed up for subscriptions, we navigated through their world using network diagrams. These incredible visuals revealed consistent connections between subscriptions and happy reviews, giving us a peek into what keeps our customers returning.

Shipping types and payment methods also got their spotlight through a heatmap, showing us the warmth of customer reviews and how they relate to how items get to our customers' doors.

Focusing on customer feedback and business strategy, Sankey diagrams helped us untangle the web of reviews, repeat purchases, subscriptions, and more. It's like deciphering the secret language of customer satisfaction.

Summing it up with waterfall charts, we saw a harmonious distribution of purchases across different items. This gave us a bird's-eye view of how our customers from other places and backgrounds shop.

In the grand journey of "Customer Insights and Strategy Optimizations," our exploration surpassed mere number crunching; it resembled an exhilarating treasure hunt. Unexpectedly, we unearthed a remarkable consistency in data across diverse regions, defying our initial expectations. These surprising and fascinating discoveries served as a potent reminder that data is not merely a collection of numbers but a source of unexpected revelations, steering us toward innovative strategies and higher customer satisfaction.

* **Link to the dashboard:**

<https://prod-uk-a.online.tableau.com/t/datpetr/views/CustomerInsightsandStrategyOptimization/Topitemcategoriesbasedonregion>

* **Link to the charts for the projects:**

<https://prod-uk-a.online.tableau.com/#/site/datpetr/workbooks/601546?:origin=card_share_link>

* **Link to the GitHub with all files:**

<https://github.com/datpetr/CUSTOMER-INSIGHTS-AND-STRATEGY-OPTIMIZATIONS>