**Checkout Flow Optimization Analysis with SQL and Tableau Project**

Boosting Online Sales: Insights Into Cart Behavior and Checkout Errors

**Retrieving Checkout Steps Information with SQL**

In this part, your assignment involves sourcing the data files that will later be imported into the Tableau workbook. You'll perform this task using MySQL, where your first step is to import the database from the **365\_checkout\_database.sql** file. Subsequently, you’ll implement the necessary queries and obtain the information for creating the dashboard graphs.

Use the tables checkout\_actionsand checkout\_carts from the **365\_checkout\_database** to retrieve a result set covering the entire period and containing the following fields:

* action\_date: the day on which the checkout activity took place
* count\_total\_carts: the count of shopping carts created each day during the specified timeframe
* count\_total\_checkout\_attempts: the count of purchase attempts each day
* count\_successful\_checkout\_attempts: the count of successful purchases each day

Creating such a result set is crucial because it provides a consolidated view of daily checkout activities over the specified period. By analyzing this data, we can better understand user behavior patterns, identify potential issues in the checkout process, and pinpoint areas for optimization to enhance the overall user experience.

To build the result set, we advise you to utilize common table expressions (CTEs) for capturing the various stages of a subscription attempt. You’ll use the WITH clause to create such common table expressions that can later be easily referenced with the help of a SELECT statement.

Save the result set as a CSV file and call it checkout\_steps.csv.

In the first three steps of the instructions, we’ll create the following CTEs:

* One that keeps all carts created
* One that stores all checkout attempts
* And one that keeps only the successful attempts.

Follow the outlined steps and substitute the question marks with the correct values and expressions when necessary.

       1. Let’s start by building the CTE that keeps all created carts. First, initialize the WITH clause and then, select all fields from the checkout\_carts table.

Note the following skeleton of this query:

???

total\_carts\_created as

(

SELECT

\*

FROM

checkout\_carts

),

        2. Now, using the table you've just set up, pull the data on all users who have created a purchase cart and have also attempted to finalize a purchase. Examine the appropriate action\_name in the relevant table to sift through and collect the required data into a new temporary result.

Note the following skeleton of this query:

total\_checkout\_attempts as

(

SELECT

tc.user\_id,

a.action\_name,

a.action\_date

FROM

??? as tc

LEFT JOIN

??? as a on a.user\_id = tc.user\_id

WHERE

a.action\_name like '%???%'

AND

a.action\_date BETWEEN '2022-07-01' AND '2023-01-31'

),

        3. Next, construct a temporary result set that captures only the successful checkout attempts. Pull the necessary data from the previous CTE, focusing on events that resulted in success.

Note the following skeleton of this query:

total\_successful\_attempts as

(

SELECT

a.user\_id,

a.action\_name,

a.action\_date

FROM

total\_checkout\_attempts as a

WHERE

???

GROUP BY a.user\_id

),

       4. With the relevant CTEs in place, it's time to implement a method to count the specific records we're targeting. We'll use additional CTEs to capture the necessary count. Begin by counting the total number of carts from the CTE containing all purchased carts daily.

Note the following skeleton of this query:

count\_total\_carts as

(

SELECT

action\_date,

???(\*) as count\_total\_carts

FROM

???

GROUP BY action\_date

),

       5. Construct a query that counts the total daily checkout attempts following the same approach.

Note the following skeleton of this query:

count\_total\_checkout\_attempts as

(

SELECT

action\_date,

???(\*) as count\_total\_checkout\_attempts

FROM

???

GROUP BY action\_date

),

       6. Now, construct a query to count only the successful daily attempts.

Note the following skeleton of this query:

count\_successful\_checkout\_attempts as

(

SELECT

action\_date,

???(\*) as count\_successful\_checkout\_attempts

FROM

???

GROUP BY action\_date

)

       7. Finally, select the pertinent data from the last three CTEs and pull out the desired columns for your checkout\_steps result set. Use the IFNULL function to substitute the null values with a value. Order the record by action\_date to obtain the information chronologically.

Note the following skeleton of this query:

SELECT

c.action\_date,

???,

IFNULL(???, 0) as count\_total\_checkout\_attempts,

IFNULL(???, 0) as count\_successful\_checkout\_attempts

FROM

count\_total\_carts c

LEFT JOIN

count\_total\_checkout\_attempts a on a.action\_date = c.action\_date

LEFT JOIN

count\_successful\_checkout\_attempts s on s.action\_date = a.action\_date

WHERE

c.action\_date BETWEEN '2022-07-01' AND '2023-01-31'

ORDER BY c.action\_date

       8. After executing the final query, save the result set as a CSV file called checkout\_steps.csv.

**Retrieving Checkout Errors Information with SQL**

Use the tables checkout\_actionsand checkout\_carts from the **365\_checkout\_database** to retrieve a new result set called checkout\_errors containing the following fields:

* user\_id – the identification number of the student attempting to checkout
* action\_date – the date of the checkout attempt
* action\_name – the text detailing the specific action performed by the user
* error\_message – the text with the received error (if any)
* device – the type of the used device (desktop or mobile)

Generating this result set is essential because it offers a comprehensive overview of checkout errors encountered by users during the defined time frame. Through its analysis, we can identify common issues students face during checkout—enabling us to make informed decisions on potential enhancements or changes to the platform.

This insight is valuable in optimizing the checkout experience and minimizing obstacles, which can directly contribute to increased user satisfaction and revenue.

Follow the steps outlined below.

Ensure to substitute the question marks with the correct values and expressions.

1. Select the pertinent columns from the checkout\_actions table that offer details about the error received post-action and the device utilized during the process. You can filter exclusively for events that resulted in errors or retrieve all actions and remove the irrelevant ones later in Tableau. Our primary interest lies in the error messages and the devices used.

Note the following skeleton of this query:

SELECT

user\_id, action\_date, action\_name, ???, ???

FROM

???

WHERE action\_date BETWEEN '2022-07-01' and '2023-01-31' and action\_name like '%checkout%'

GROUP BY user\_id

ORDER BY action\_date

1. After executing the query, save the result set as a CSV file called checkout\_errors.csv.

**Creating the Charts in Tableau: Monthly Checkout Success Rate**

To construct the desired story, begin by creating individual sheets in Tableau. Next, merge these sheets into dashboards and position the dashboards on distinct story points. This approach ensures you have a comprehensive, interactive visualization tool to aid your data analysis.

Begin by utilizing the **checkout\_steps** data source to craft a dual graph illustrating the monthly checkout success rate. This graph should display the total number of checkout attempts and the percentage of successful ones. Proceed with the steps provided below to accomplish this.

1. Import the checkout\_steps CSV file as a data source in Tableau Public.
2. Open a new sheet and start building the first graph by placing the Action Date field into Columns and adjusting the date format from year to month. Ensure the values are set to Discrete.
3. Drag the Count Total Checkout Attempts field into Rows. Show the mark labels and modify the graph as a bar chart. For better visibility, switch to an Entire View.
4. Next, compute a new field to determine the monthly checkout success rate. Navigate to Create Calculated Field, label it Checkout Success Rate, and input the appropriate formula in the provided space. The success rate is calculated by dividing the sum of the successful checkout attempts by the total checkout attempts and converting the result into percentages. Use the SUM function to identify the two categories of attempts. After the field is made, place it into Rows to obtain the second graph.
5. Adjust the axis and pane settings of the second graph to display values as percentages using the Format tool. Change the visualization to a line chart. To merge the two graphs, use the Dual Axis feature after right-clicking on the Checkout Success Rate field.
6. Go to Edit Axis and fix the range to start from 0 and end to 1 for better visibility.
7. Add a time range filter by placing the Action Date field into the filter section and displaying it with Show Filter. Set the graph title as Monthly Checkout Success Rate and finalize the sheet.

**Creating the Charts in Tableau: Monthly Cart Abandonment Rate**

Now that you’re ready with the first visualization, use the **checkout\_steps** table to create another dual graph about the monthly number of purchase carts and the relevant abandonment rate.

Proceed with the steps provided below to create the necessary visualization.

1. Create a new sheet and drag the Action Date field into Columns. Adjust the date format from year to months. Ensure the values are set to discrete.
2. Drag the Count Total Carts field into Rows. Show the mark labels and modify the graph as a bar chart. For better visibility, switch to Entire View.
3. Next, create a new calculated field to determine the monthly cart abandonment rate. Navigate to Create Calculated Field, label it Abandonment Rate, and input the appropriate formula in the provided space. The cart abandonment rate typically refers to the portion of users who added items to their cart but did not proceed to the checkout stage—they left without attempting to checkout. To calculate this rate, subtract the total number of checkout attempts from the overall cart count, and divide the result by the total carts. Format the cart abandonment rate as a percentage.

 Cart abandonment rate=(Total carts − Checkout attempts)Total cartsCart abandonment rate=(Total carts − Checkout attempts)Total carts

In Tableau, you'll need to subtract the value in the Count Total Checkout Attempts field from the Count Total Carts field. Both should be within the SUM function. Then, divide this result by the SUM of the Count Total Carts field.

After the field is made, place it into Rows to obtain the second graph.

      4. Adjust the axis and pane settings of the second graph to display values as percentages using Format. Change the visualization to a line chart. To merge the two graphs, use the Dual Axis feature by right-clicking on the Cart Abandonment Rate field.  
      5. You can extend the range of the y-axis to 100% by right-clicking on it and using the Edit Axis option. Go to Edit Axis and fix the range to start from 0 and end to 1.  
      6. Add a time range filter by placing the Action Date field into the filter section and displaying it with Show Filter. Set the graph title as Monthly Cart Abandonment Rate and finalize the sheet.

**Creating the Charts in Tableau: Error Messages**

The third visualization for the story shows the most frequent error messages encountered during failed checkout attempts. Create a horizontal bar chart showing the most common errors.

Proceed with the steps provided below to create the necessary visualization.

1. Import the checkout\_errors CSV file as a data source in Tableau Public.
2. Drag the Error Message field into Rows. Create a new calculated field and place the checkout\_errors.csv generated field into the empty area, naming it Attempts. Then, drag the Attempts field into the Columns shelf.
3. If you've chosen to export all checkout attempts—including the successful ones—a Null title should now appear at the top of the rankings. This represents all attempts that didn't produce an error. For the current analysis, exclude these attempts. Order the results in descending fashion using Tableau's quick sort feature and select Entire View for a more precise display.
4. Place the Device fields into the Color property in the Marks card to differentiate each bar into mobile and desktop segments. This will help you quickly identify correlations between the errors and the devices used. Show mark labels to see the values behind the stacked bars.
5. Finally, add a time range and device filters and title the graph Checkout Error Messages. You can exclude the 10 least frequent errors from the chart since they rarely occur and don't significantly impact the visualization.

**Creating the Charts in Tableau: Device Distribution**

For your upcoming task, create a horizontal stacked bar chart that categorizes device usage into percentages for mobile and desktop. This will let you quickly determine which device is more commonly used.

1. Open a new sheet and place the Device field into Rows.
2. Next, drag the Attempts calculated field into Columns and show the mark labels to see the number of attempts made with each type of device. Enlarge the graph for better visibility and modify it as a horizontal stacked bar chart.
3. Click on the Attempts field, go to Quick Table Calculation, and choose Percent of Total to visualize the separation proportionally.
4. Name the graph Desktop vs Mobile and complete the sheet.

**Creating the Tableau Story**

Well done on completing all the necessary graphs! Now, your subsequent task is to distribute these graphs over three story points.

* **Story Point 1** includes the dual graph of monthly checkout attempts and the relevant success rates.
* **Story Point 2** comprises the dual graph about the monthly number of checkout carts and the relevant abandonment rates.
* **Story Point 3** encompasses two horizontal stacked bar charts: one with the most common error messages and one with the proportions of devices used for checkout.

1. First, create a new story sheet from the relevant button in Tableau. Then, adjust its dimensions to a fixed size of a generic desktop. Now, you’re ready to drag the appropriate sheets into the view.
2. Drag the worksheet with the monthly checkout success rate on the story’s first page and label it with the relevant name.
3. Add a new story point and drag the worksheet with the monthly cart abandonment rate. Insert an appropriate caption as a name.
4. The third story point will consist of two graphs. Therefore, start by creating a dashboard that will act as a platform to display the two visualizations. Create a new dashboard and fix its size to fit the story. Next, position the error messages graph at the top and set the device distribution chart underneath it. Shrink the size of the device bar chart to allow more room for the graph on top.
5. Modify the time range filter of the error messages graph to affect the device graph. Navigate to each filter, select Apply to Worksheets, and choose All Using This Data Source. Check if the filter is working correctly.
6. Create another blank story point and drag the dashboard into its view. Name it Error Messages and Devices.
7. **Formatting the Story**
8. Now that you've completed the story-based dashboard, it's time to fine-tune its appearance. Modify the graph colors, tweak the dimensions, and add fitting titles.
9. 1. Assign the title Checkout Flow Optimization Dashboard to the story and align the text to the center. Go to Format and change the shading of the story to dark grey (HTML color code #545c69), and the navigator’s shading must be in lighter grey (HTML #e6e7e9). The title’s color can now be switched to white. Enlarge the navigator box dimensions to ensure the labels display on one line.
10. 2. Next, modify the colors of the individual charts. Begin with the Checkout Success Rate chart. Navigate to the worksheet's Marks card and select the Color property. The color of the line chart must have an HTML #293343; the bars must be in #3dafb8.
11. 3. Go to the graph on Story Point 2 and modify its colors from the same menu. Color the line chart in #550000 and the bars in #8cdae3.
12. 4. The third page of the story-based dashboard is about the errors and the devices used for checkout. Color the desktop bars in #293343 and the mobiles in #3a9ea7. In the error messages worksheet, remove the subtitle to enhance clarity. For the same reason, you can change the bar’s mark label’s color from black to white.
13. 5. Once you complete these steps, you can save your story-based dashboard and draw crucial insights about the checkout process.

**Interpreting the Results**

Now that you’ve successfully created your Tableau dashboard, you’re ready to gain valuable insights from it and propose areas of improvement. Based on the findings gathered from the Tableau dashboard, provide a comprehensive **analysis report** including the **current state of affairs**, **business objective**, your chosen **hypothesis**, and suggested **actionable insights**. Justify your selections based on data interpretations and potential impacts on the overall user experience.

Begin by examining the story pages and identifying trends.

* Are there months that underperform or outperform others significantly?
* Can you discern a correlation between the error messages received and the devices on which they occurred?
* What suggestions can you make to enhance the overall checkout process?

**Current State of Affairs**

Examine each story page sequentially to outline the checkout flow’s current state. Identify the months with the highest checkout success rates, those with the highest cart abandonment rates, the predominant error message, and noticeable patterns related to device usage.

**Business Objective**

The business goal concerning the checkout flow centers on refining and optimizing the process for a better user experience. From a business standpoint, this increases sales, boosting profits.

**Hypothesis**

The findings gathered from the dashboard allow for multiple hypotheses addressing the visible challenges in the platform's checkout flow. Our focus should be on identifying the most prominent issue tied to the checkout errors and seeking a straightforward remedy. Consider a simple solution, especially given the prevalence of the most common error message on mobile devices compared to desktops.

**Actionable Insights**

Given the findings we've gathered, we can now propose various actionable strategies to address users’ challenges with their debit and credit card issues. These may differ based on the specific improvement areas we focus on. Consider suggestions grounded in your formulated hypothesis, but also delve into alternative solutions that address payment challenges identified from your observations.

When seeking actionable solutions to improve the checkout process, consider the following guiding insights and pose pertinent questions in four specific domains:

1. **Webpage Interface Enhancements**

* How easy is it for users to navigate the checkout page?
* Is the process intuitive and straightforward, or does it require unnecessary steps?
* Are the payment information input fields large enough?
* How does the checkout interface adapt to different device sizes and screen orientations?

1. **Demographics**

* Are there elements of the checkout process that might not be as effective for specific demographic groups?
* Are all parts of the checkout process translated appropriately for users from different regions?

1. **Payment Alternatives**

* Does the checkout process offer multiple payment methods to cater to a broader range of users?
* Are there emerging payment methods in certain parts of the world that still need to be integrated?

1. **Real-time card validation**

* Which error types occur most frequently?
* Would the process improve if users were alerted to issues with their bank cards while entering the necessary details?

**Quiz**

**Question 1:**

In the context of our project on checkout flow optimization, understanding monthly performance is crucial to pinpointing areas of improvement and potential anomalies. Leveraging the story you've constructed, which month—over the entire analysis period—stands out as the most successful in terms of checkout attempts?

November 2022

**Question 2:**

Working on such an analysis, it’s essential to identify the peaks and trenches in our monthly performance. By delving into the story you've assembled, can you determine which month—throughout the entire analysis duration—witnessed the lowest number of checkout attempts, and what was that exact figure?

October 2022, attempt ---173

**Question 3:**

As we aim to refine the checkout flow and understand user behavior, pinpointing our peak moments can provide significant insights. From your constructed story, can you discern which of the listed months registered the highest count of purchase carts?

January 2023

**Question 4:**

Identifying cart abandonment patterns is crucial to effectively strategizing improvements in our checkout flow. Based on the narrative presented in your story, can you identify the two months that witnessed the most significant cart abandonment rates?

August 2022 and October 2022August 2022 and November 2023July 2022 and December 2022October 2022 and January 2023

Novembere 2022 – Sep 2022 – Jan 2023

**Question 5:**

Understanding device-related error trends is crucial to finding and rectifying user issues during checkout. Can you determine the error message that most frequently appeared on desktop devices during September 2022?

Year field is required

**Question 6:**

To ensure the accuracy of our actions moving forward, we must verify our insights against the data presented in the story. Referring to the narrative in your story, can you identify which of the following findings aligns with the information from the analysis?

The checkout success rate in September 2022 is approximately 40%.The cart abandonment rate in September 2022 is lower than in December 2022.The error message, Your Card Has Insufficient Funds, is the third most common overall.Your Card was Declined is the third most common error received on the desktop.

**Question 7:**

Opportunity sizing refers to estimating the potential value or impact of a particular opportunity or solution. In the context of business strategy and decision-making, it helps quantify the potential benefit of an initiative, giving a clearer picture of the stakes at hand.

For our Checkout Flow Optimization project, let's delve into opportunity sizing:

* In January, the checkout success rate was 34% out of 360 attempts.
* Each successful purchase generates approximately $30 in revenue.
* We aim to improve the checkout success rate to 40% in February.

For this exercise, assume that the number of checkout attempts in February will remain consistent with January (360 attempts). Given this assumption, calculate the opportunity size. Compared to January's earnings, how much additional revenue could the company earn in February if we achieve the targeted 40% checkout success rate?

$547$760$567$648