

Data Analysis on Bike Store

Introduction: Welcome to the dynamic world of a Bike Store in USA , where the thrill of cycling meets the precision of data. In this exciting chapter of our journey, we delve into the intricate realm of data analysis to understand and optimize every aspect of our bike store's operations. Join us on this expedition as we navigate the data trail, unlocking insights that propel us forward in delivering an unparalleled cycling experience.

Data-Driven Pedals: our commitment to excellence extends beyond the workshop floor and into the realm of data analytics. By harnessing the power of data, we aim to fine-tune our operations, enhance customer experiences, and elevate our understanding of the ever-evolving cycling landscape.

Exploring the Analytics Landscape: Just as cyclists explore diverse terrains, we explore the vast landscape of data analytics. From sales trends to customer behavior, inventory management to marketing effectiveness, our data analysis project is a comprehensive expedition that aims to uncover patterns, optimize strategies, and ensure that every gear shift in our store is backed by informed decisions.

Your Ride, Your Data: In this data-driven era, we understand the importance of personalization. Our data analysis project isn't just about numbers; it's about tailoring the cycling experience to your preferences. By analyzing your interactions, purchases, and feedback, we strive to create a more customized and enjoyable journey for every cyclist who steps into the Bike Store.

Joining Tables Using MySQL

Introduction: In the realm of data analysis for our bike sales project, SQL joins emerge as the crucial gears that seamlessly connect disparate datasets, enabling us to extract valuable insights from the intricate tapestry of information. This brief explores how SQL joins elevate our capacity to understand, analyze, and enhance the performance of our bike sales at [Your Bike Store].

Understanding the Terrain: Our bike sales data spans various dimensions — from customer information and product details to transaction records and inventory status. SQL joins act as the navigational tool, allowing us to traverse this multidimensional landscape with precision and clarity.

Accelerating Analysis:

- **Efficiency:** SQL joins streamline the process, allowing us to retrieve relevant information from multiple tables in a single query.
- **Accuracy:** By establishing relationships between tables, we ensure that our analyses are based on comprehensive and interconnected datasets.
- **Insight Generation:** SQL joins empower us to unveil intricate patterns, correlations, and trends that drive informed decision-making in product placement, marketing strategies, and inventory management.

MySQL Query:

```
SELECT
    ord.order_id,
    CONCAT(cus.first_name," ",cus.last_name) AS "customer",
    cus.city,
    cus.state,
    SUM(ite.quantity) AS "total_units",
    SUM(ite.quantity * ite.list_price) AS "revenue",
    pro.product_name,
    cat.category_name,
    sto.store_name,
    CONCAT(sta.first_name , " " , sta.last_name) AS "Sales_rep"
FROM    orders ord
    JOIN customers cus
        ON ord.customer_id = cus.customer_id
    JOIN order_items ite
        ON ord.order_id = ite.order_id
    JOIN products pro
        ON ite.product_id = pro.product_id
    JOIN categories cat
        ON pro.category_id = cat.category_id
    JOIN stores sto
        ON ord.store_id = sto.store_id
    JOIN staffs sta
        ON ord.staff_id = sta.staff_id

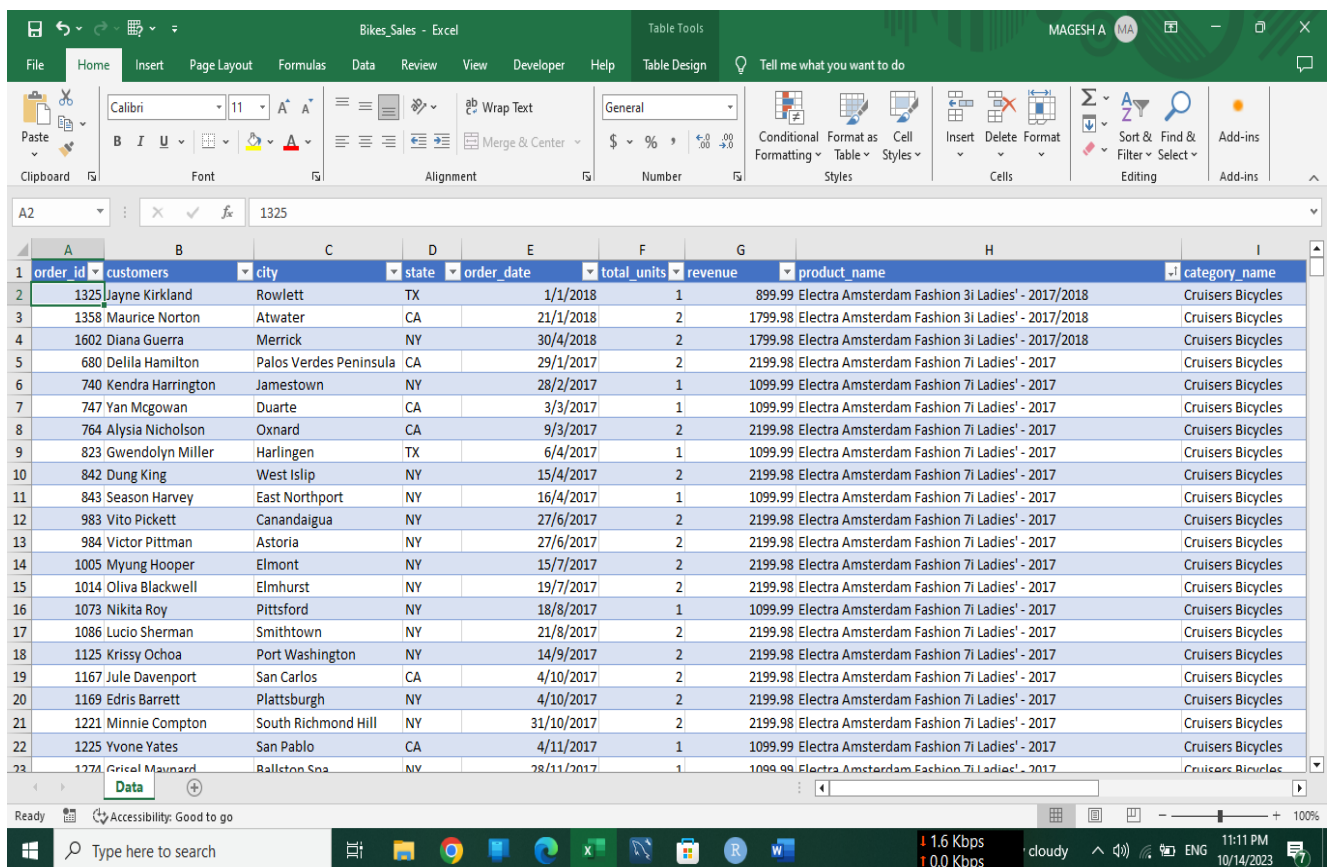
GROUP BY
    ord.order_id,
    CONCAT(cus.first_name , " " , cus.last_name),
    cus.city,
    cus.state,
    ord.order_date,
    pro.product_name,
    cat.category_name,
    CONCAT(sta.first_name , " " , sta.last_name)
```

Data cleaning

Introduction: In the world of data analysis, the importance of clean, well-organized data cannot be overstated. Excel, a powerhouse in spreadsheet software, provides robust tools for data cleaning, allowing analysts to transform raw datasets into refined gems. This brief explores the essential techniques and tools within Excel for effective data cleaning.

Identifying Data Issues: Before diving into the cleaning process, it's crucial to identify common data issues such as missing values, duplicates, inconsistencies, and formatting errors. Excel's data cleaning features are designed to address these challenges and enhance the overall quality of your dataset.

Dataset -> [click here](#)



order_id	customers	city	state	order_date	total_units	revenue	product_name	category_name
1325	Jayne Kirkland	Rowlett	TX	1/1/2018	1	899.99	Electra Amsterdam Fashion 3i Ladies' - 2017/2018	Cruisers Bicycles
1358	Maurice Norton	Atwater	CA	21/1/2018	2	1799.98	Electra Amsterdam Fashion 3i Ladies' - 2017/2018	Cruisers Bicycles
1602	Diana Guerra	Merrick	NY	30/4/2018	2	1799.98	Electra Amsterdam Fashion 3i Ladies' - 2017/2018	Cruisers Bicycles
680	Delila Hamilton	Palos Verdes Peninsula	CA	29/1/2017	2	2199.98	Electra Amsterdam Fashion 7i Ladies' - 2017	Cruisers Bicycles
740	Kendra Harrington	Jamestown	NY	28/2/2017	1	1099.99	Electra Amsterdam Fashion 7i Ladies' - 2017	Cruisers Bicycles
747	Yan McGowan	Duarte	CA	3/3/2017	1	1099.99	Electra Amsterdam Fashion 7i Ladies' - 2017	Cruisers Bicycles
764	Alysia Nicholson	Oxnard	CA	9/3/2017	2	2199.98	Electra Amsterdam Fashion 7i Ladies' - 2017	Cruisers Bicycles
823	Gwendolyn Miller	Harlingen	TX	6/4/2017	1	1099.99	Electra Amsterdam Fashion 7i Ladies' - 2017	Cruisers Bicycles
842	Dung King	West Islip	NY	15/4/2017	2	2199.98	Electra Amsterdam Fashion 7i Ladies' - 2017	Cruisers Bicycles
843	Season Harvey	East Northport	NY	16/4/2017	1	1099.99	Electra Amsterdam Fashion 7i Ladies' - 2017	Cruisers Bicycles
983	Vito Pickett	Canandaigua	NY	27/6/2017	2	2199.98	Electra Amsterdam Fashion 7i Ladies' - 2017	Cruisers Bicycles
984	Victor Pittman	Astoria	NY	27/6/2017	2	2199.98	Electra Amsterdam Fashion 7i Ladies' - 2017	Cruisers Bicycles
1005	Myung Hooper	Elmont	NY	15/7/2017	2	2199.98	Electra Amsterdam Fashion 7i Ladies' - 2017	Cruisers Bicycles
1014	Oliva Blackwell	Elmhurst	NY	19/7/2017	2	2199.98	Electra Amsterdam Fashion 7i Ladies' - 2017	Cruisers Bicycles
1073	Nikita Roy	Pittsford	NY	18/8/2017	1	1099.99	Electra Amsterdam Fashion 7i Ladies' - 2017	Cruisers Bicycles
1086	Lucio Sherman	Smithtown	NY	21/8/2017	2	2199.98	Electra Amsterdam Fashion 7i Ladies' - 2017	Cruisers Bicycles
1125	Krissy Ochoa	Port Washington	NY	14/9/2017	2	2199.98	Electra Amsterdam Fashion 7i Ladies' - 2017	Cruisers Bicycles
1167	Jule Davenport	San Carlos	CA	4/10/2017	2	2199.98	Electra Amsterdam Fashion 7i Ladies' - 2017	Cruisers Bicycles
1169	Edris Barrett	Plattsburgh	NY	4/10/2017	2	2199.98	Electra Amsterdam Fashion 7i Ladies' - 2017	Cruisers Bicycles
1221	Minnie Compton	South Richmond Hill	NY	31/10/2017	2	2199.98	Electra Amsterdam Fashion 7i Ladies' - 2017	Cruisers Bicycles
1225	Yvonne Yates	San Pablo	CA	4/11/2017	1	1099.99	Electra Amsterdam Fashion 7i Ladies' - 2017	Cruisers Bicycles
1274	Grisel Maunard	Ballston Spa	NY	28/11/2017	1	1099.99	Electra Amsterdam Fashion 7i Ladies' - 2017	Cruisers Bicycles

Analyzing using R

Installing Packages

```
install.packages("dplyr")
install.packages("tidyverse")
install.packages("tidyr")
install.packages("ggplot2")
install.packages("rmarkdown")
install.packages('tinytex')
library(dplyr)
library(tidyverse)
library(tidyr)
library(ggplot2)
library(rmarkdown)
library(tinytex)
```

Reading input

```
Bikes_Sales <-read_excel("Bikes_Sales.xlsx")
```

Column names

```
colnames(Bikes_Sales)
```

```
## [1] "order_id"      "customers"     "city"          "state"  
## [5] "order_date"    "total_units"   "revenue"       "product_name"  
## [9] "category_name" "brand_name"    "store_name"    "sales_rep"
```

head()

```
head(Bikes_Sales)
```

```
## # A tibble: 6 x 12
##   order_id customers      city state order_date      total_units revenue
##   <dbl> <chr>          <chr> <chr> <dtm>          <dbl>   <dbl>
## 1    1325 Jayne Kirkland Rowl... TX    2018-01-01 00:00:00         1    900.
## 2    1358 Maurice Norton  Atwa... CA    2018-01-21 00:00:00         2   1800.
## 3    1602 Diana Guerra    Merr... NY    2018-04-30 00:00:00         2   1800.
## 4     680 Delila Hamilton Palo... CA    2017-01-29 00:00:00         2   2200.
## 5     740 Kendra Harrington Jame... NY    2017-02-28 00:00:00         1   1100.
## 6     747 Yan McGowan     Duar... CA    2017-03-03 00:00:00         1   1100.
## # i 5 more variables: product_name <chr>, category_name <chr>,
## #   brand_name <chr>, store_name <chr>, sales_rep <chr>
```

str()

```
str(Bikes_Sales)
```

```
## tibble [4,722 × 12] (S3: tbl_df/tbl/data.frame)
## $ order_id      : num [1:4722] 1325 1358 1602 680 740 ...
## $ customers     : chr [1:4722] "Jayne Kirkland" "Maurice Norton" "Diana Guerra" "Delila Hamilton" ...
## $ city          : chr [1:4722] "Rowlett" "Atwater" "Merrick" "Palos Verdes Peninsula" ...
## $ state         : chr [1:4722] "TX" "CA" "NY" "CA" ...
## $ order_date    : POSIXct[1:4722], format: "2018-01-01" "2018-01-21" ...
## $ total_units   : num [1:4722] 1 2 2 2 1 1 2 1 2 1 ...
## $ revenue       : num [1:4722] 900 1800 1800 2200 1100 ...
## $ product_name  : chr [1:4722] "Electra Amsterdam Fashion 3i Ladies' - 2017/2018" "Electra Amsterdam Fashion 3i Ladies' - 2017/2018" "Electra Amsterdam Fashion 3i Ladies' - 2017/2018" "Electra Amsterdam Fashion 7i Ladies' - 2017" ...
## $ category_name: chr [1:4722] "Cruisers Bicycles" "Cruisers Bicycles" "Cruisers Bicycles" "Cruisers Bicycles" ...
## $ brand_name    : chr [1:4722] "Electra" "Electra" "Electra" "Electra" ...
## $ store_name    : chr [1:4722] "Rowlett Bikes" "Santa Cruz Bikes" "Baldwin Bikes" "Santa Cruz Bikes" ...
## $ sales_rep     : chr [1:4722] "Kali Vargas" "Mireya Copeland" "Marcelene Boyer" "Genna Serrano" ...
```

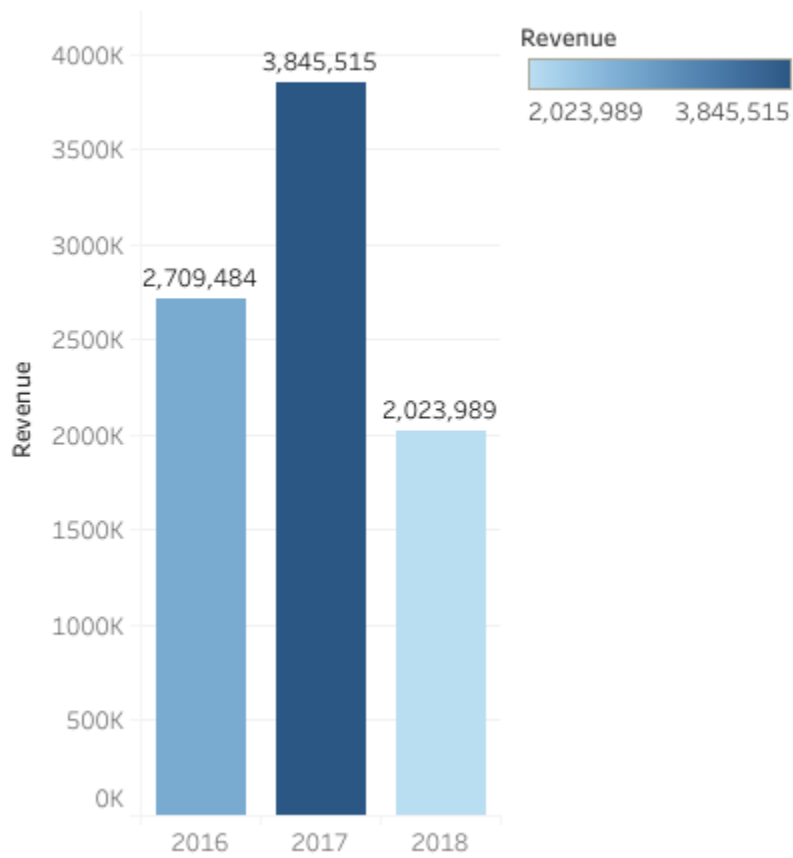

summary()

```
summary(Bikes_Sales)
```

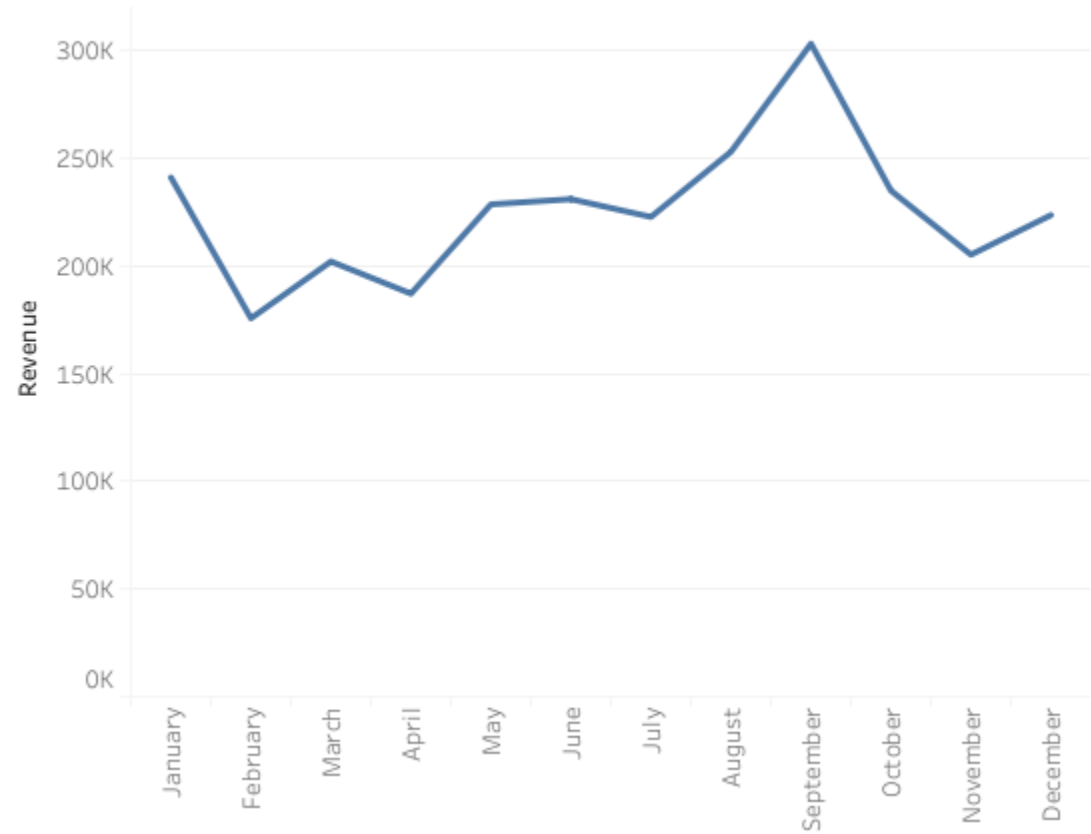
```
##      order_id      customers      city      state
## Min.   : 1.0    Length:4722    Length:4722    Length:4722
## 1st Qu.: 423.2  Class :character  Class :character  Class :character
## Median : 828.5  Mode  :character  Mode  :character  Mode  :character
## Mean   : 821.3
## 3rd Qu.:1226.0
## Max.   :1615.0
##      order_date      total_units      revenue
## Min.   :2016-01-01 00:00:00.00  Min.   :1.000  Min.   : 89.99
## 1st Qu.:2016-09-06 12:00:00.00  1st Qu.:1.000  1st Qu.: 539.98
## Median :2017-04-09 00:00:00.00  Median :1.000  Median : 939.98
## Mean   :2017-04-04 22:18:26.99  Mean   :1.499  Mean   :1816.81
## 3rd Qu.:2017-11-04 00:00:00.00  3rd Qu.:2.000  3rd Qu.:1999.98
## Max.   :2018-12-28 00:00:00.00  Max.   :2.000  Max.   :23999.98
## product_name      category_name      brand_name      store_name
## Length:4722      Length:4722      Length:4722      Length:4722
## Class :character  Class :character  Class :character  Class :character
## Mode  :character  Mode  :character  Mode  :character  Mode  :character
##
##
##
##      sales_rep
## Length:4722
## Class :character
## Mode  :character
##
##
##
```

Data Viz

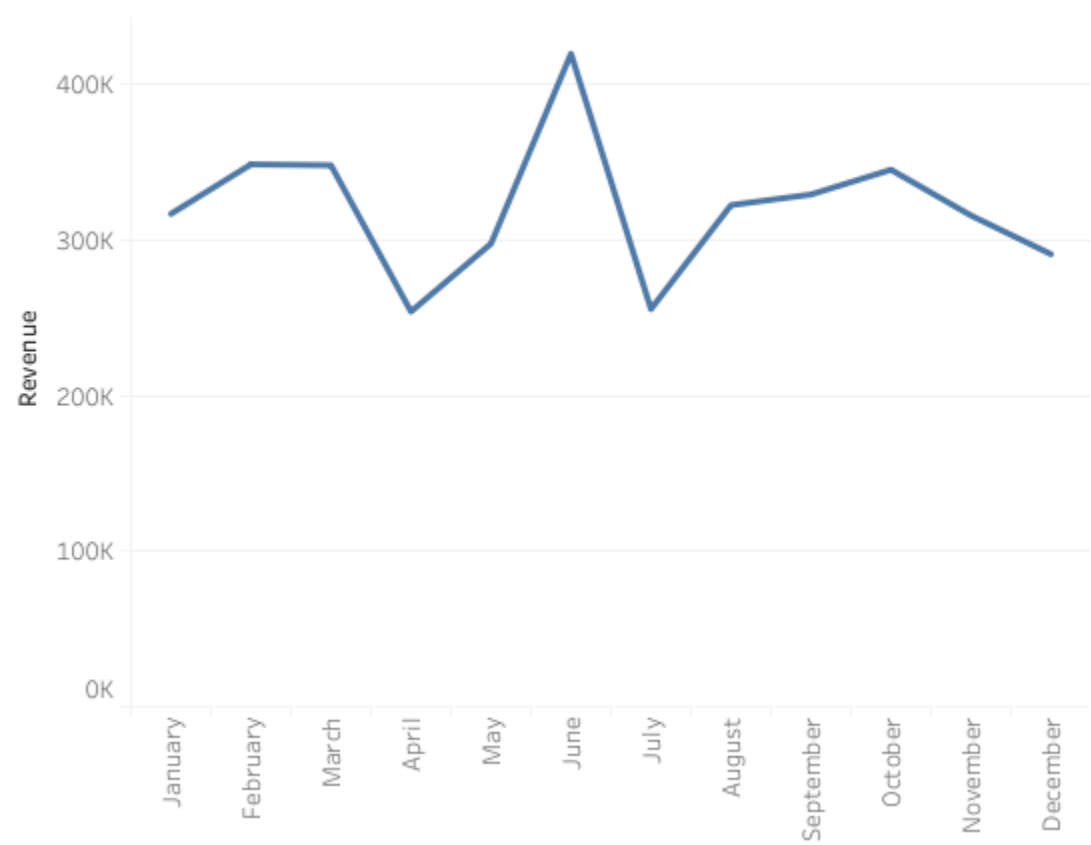
Total Revenue



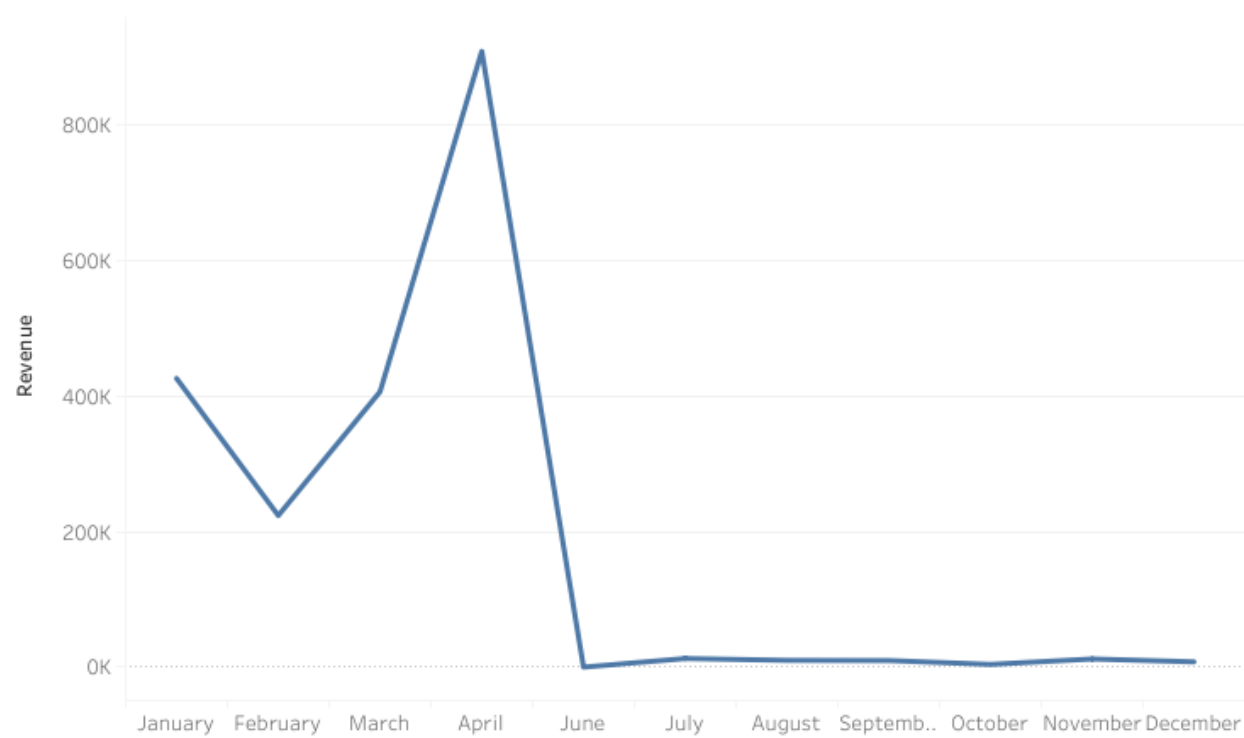
Revenue per month - 2016



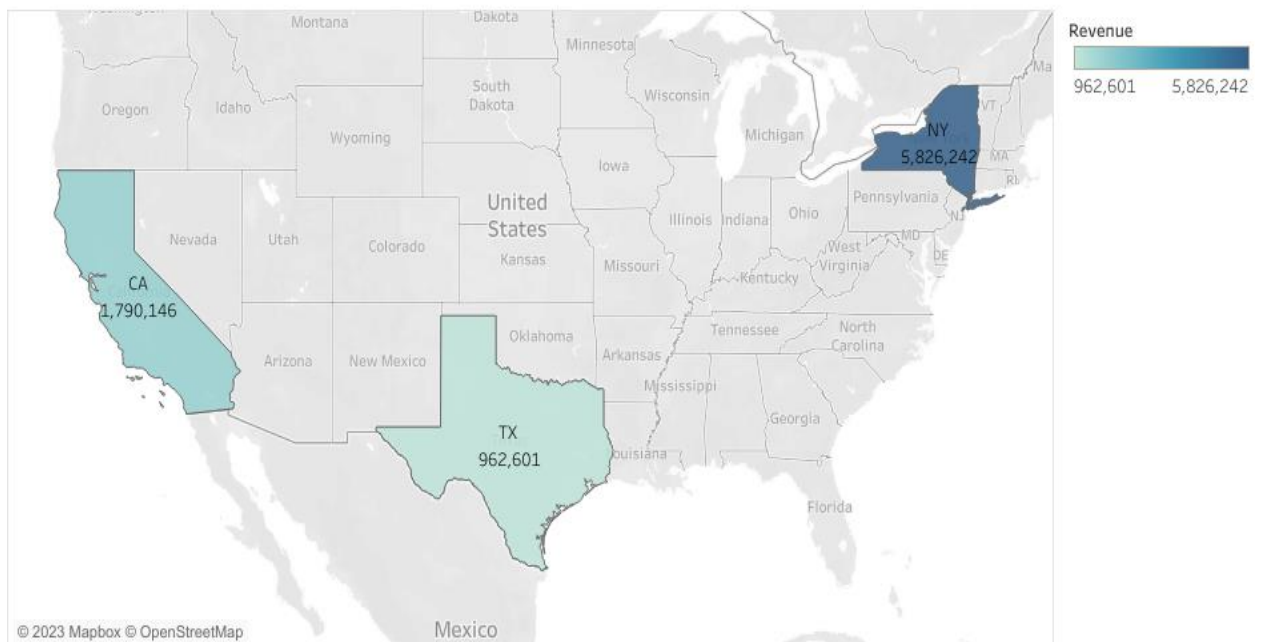
Revenue per month - 2017



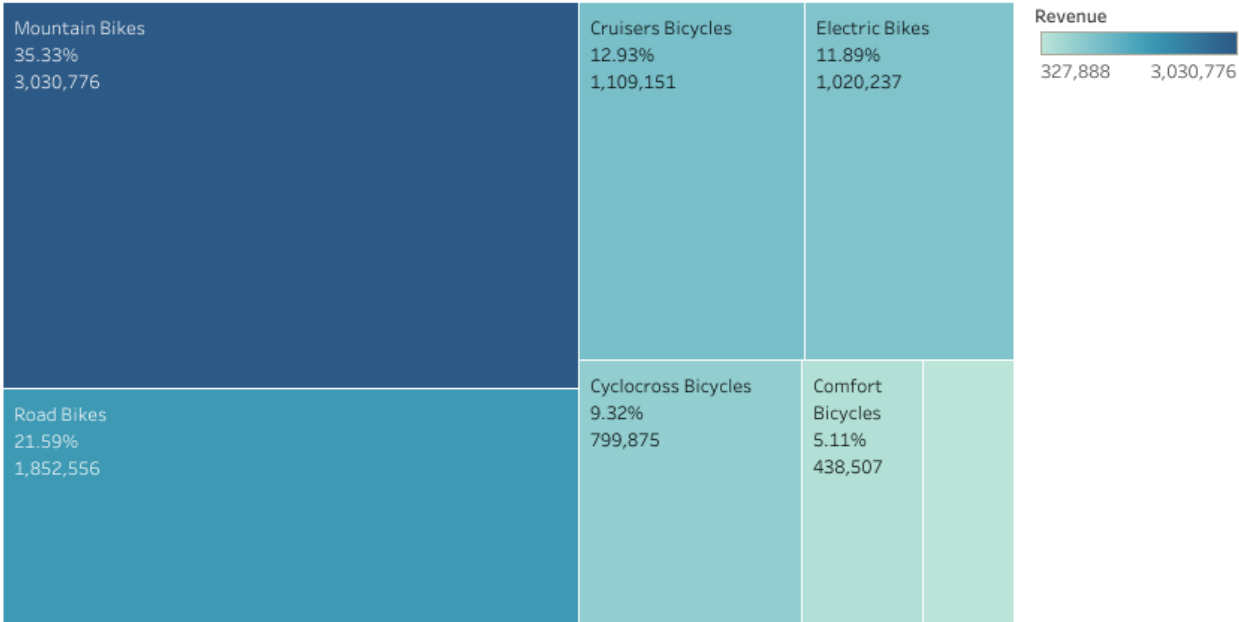
Revenue per month - 2018



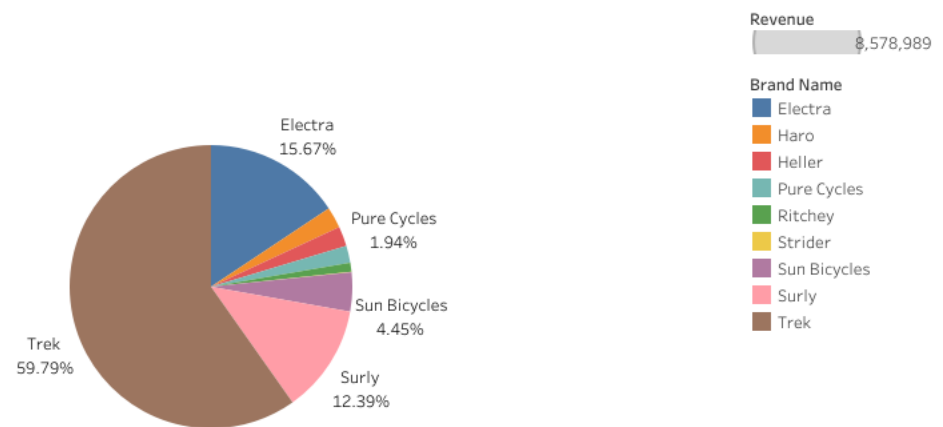
Revenue by State



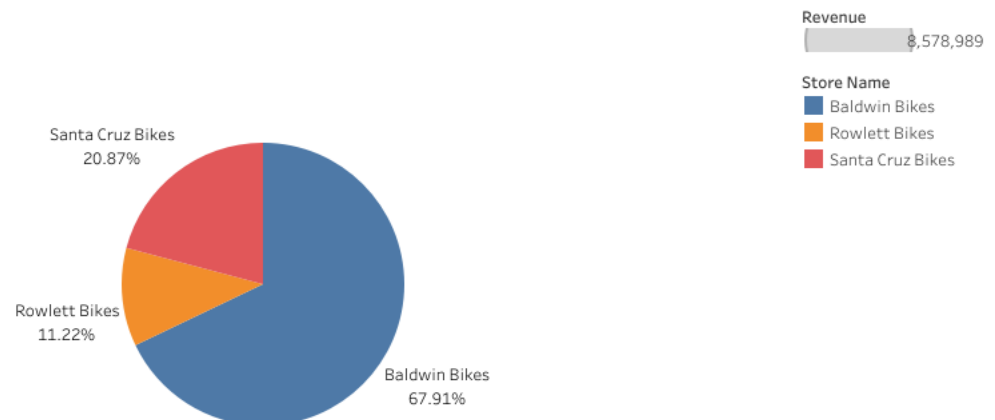
Revenue by category



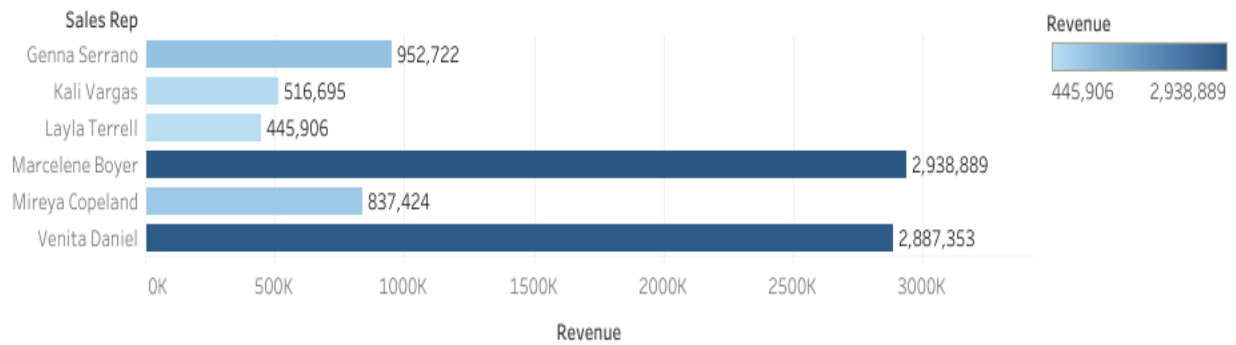
Revenue per Brand



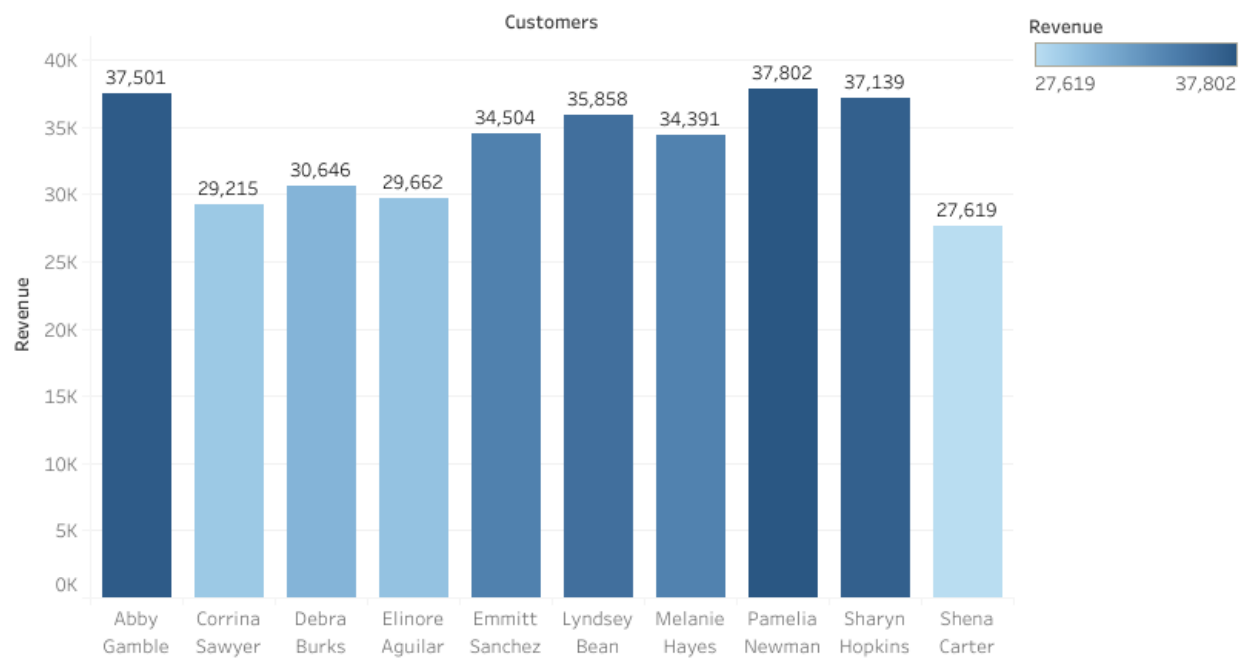
Revenue per Store



Top Sales Rep



Top Customers



Data Driven Decision Making:

- The Electric vehicles category is low and contributing major part in the revenue So, Increase the category of E – Vehicles.
- The number of stores of Rowlett Bikes is low and ratio of revenue of this store is marginally higher than other store So, Increase the number of stores.
- The number of stores in the California city is high and the revenue ratio is low So, reduce the number of the stores in California.
- Customer's are preferring Electra – Brand the more than the other brand but their revenue ratio is considerably low So, launch new products from Electra.
- The number of times visiting the store by a customer is pretty low So, by giving give away and gift to the selected repeated customers will increase the revenue.
- By increasing the salary of the top sales representative create a boost among the sales rep.
- The revenue graph is not uniform , company need to work on the customer feedback and give what customer exactly needed.

