Question 4 :

a. Tell your experience in past company or campus, if you are in charge in Data Analyst role, how is your day to day is look like and what expectations do you have for Vidio as Data Analyst role.
b. Tell what study case that you already solved using data in your previous work. If the data is confidential, just blur the number and the explicit information regarding the company.

---

a. I had experienced as a programmer in past company, so I have a very solid understanding of databases, including database design, advanced query, PL/SQL, advanced SQL for data engineers and well-versed in numerous programming languages. I also have experience as an Assistant Lecturer of Basic Computation Practicum when I was in college that covers mathematical analysis (calculus and statistics) using Matlab software.

   Data analyst and data science is a slice of the fields that be my passions and are best suited to my most prominent skills, i.e. I like database programming the most; mathematics has been my favorite subject since elementary school, and I got the best score in Statistics and Calculus when I was in college; I have previous work experience as a programmer. So, mathematics, database programming, and programming languages are very close to me, and I have strong statistical & analytical thinking and strong data visualization skills.

   I completed and earned the Google Data Analytics Professional Certificate – rigorous, hands-on program that covers the entire scope of the data analysis process.
   Also recently completed "Introduction to Statistics" by Stanford University and "Python for Data Science, AI & Development" by IBM.
   Credential URL of all certificates that I earned and short description about the courses could be seen in Curriculum Vitae.
   Now, I am in the process of completing the IBM Data Science Professional Certificate.

   I have excellent understanding and proficiency of platforms for effective data analyses, including SQL, Excel, Tableau, and Python. Confidence in explore complex data into actionable and clear insights; build rich, insightful, and beautiful data visualization and dashboards.

   Expectations that I have for Vidio as Data Analyst role :

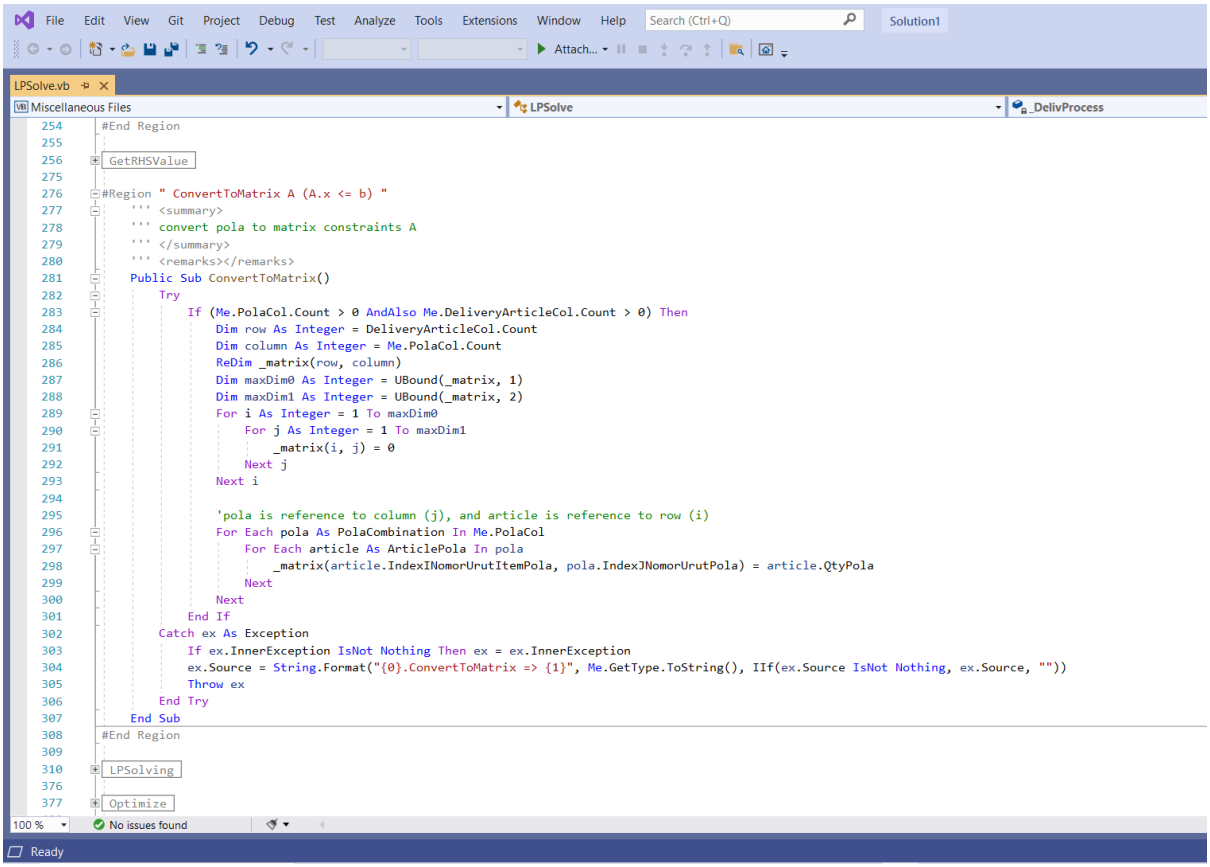
   I'm looking for the opportunity to expand my learning, put to use my skills and experience in data analysis, work in a team that is inclusive and a positive workplace environment that helps propel my growth beyond the job descriptions.

b. I have experience working on projects using mathematical modeling using Linear Programming and Object Oriented Programming for optimizing available stock in warehouse to container allocation.

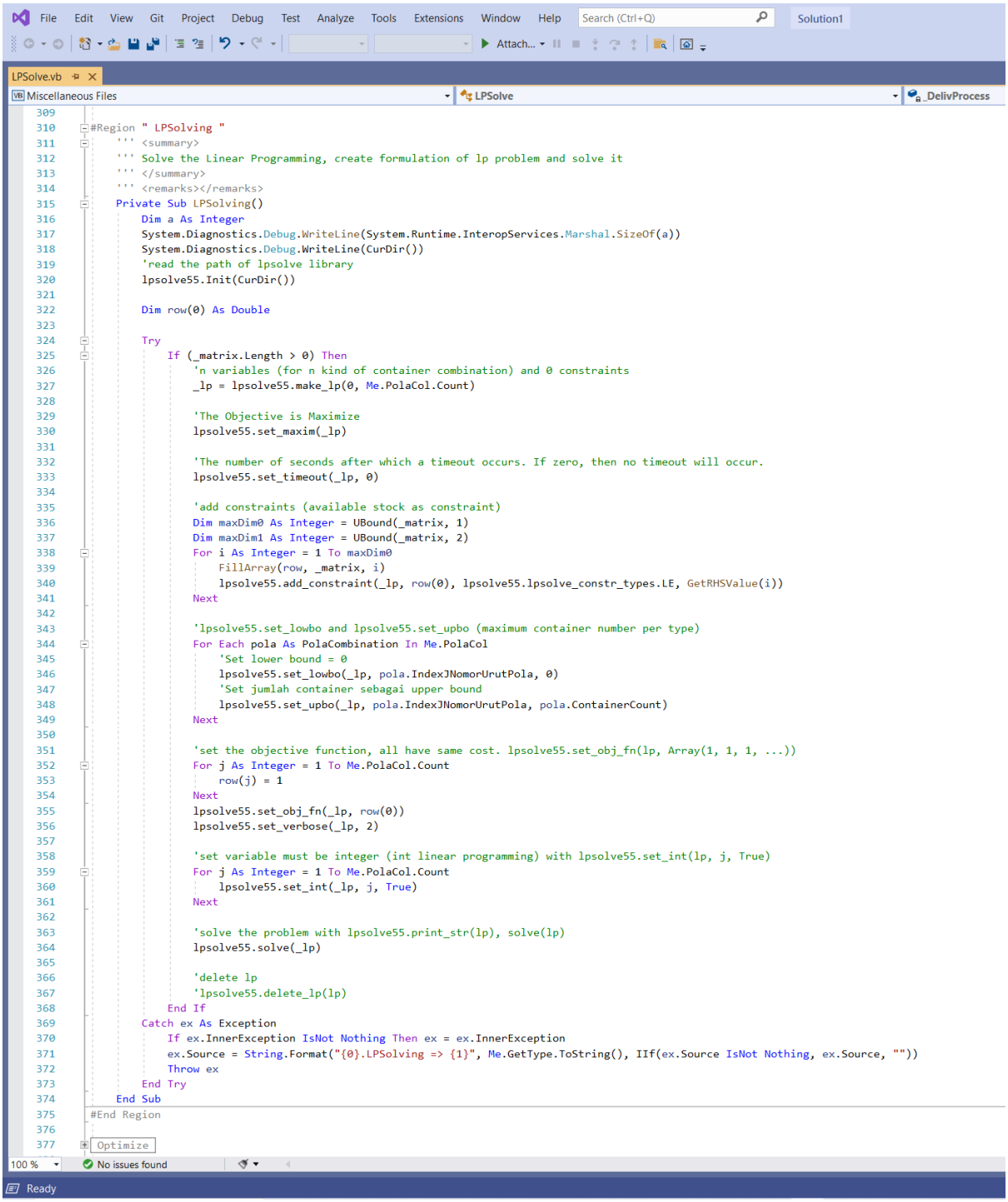
1) Project Name : **Cargo Ready Container System**
   Client : PT Pindo Deli Pulp & Paper Mills
   Role : Programmer
   Software and Platform : SQL Server, IIS, Visual Studio, LPSolver, SAP, E-Draw, Innovasys Help
   Skill sets used : VB.NET, ASP.NET, SQL, SAP, UML, Linear Programming
   → Developed system for optimizing available stock in warehouse to container allocation using Linear Programming methodology and Object Oriented Programming.
   Optimizing solution for how many containers can be ready in delivery data (each container consists of many materials with different quantities) with stock availability.
   Modeling the problems mathematically (using matrix) so that can be solved with linear programming.

   Here are some snippets of the program code :

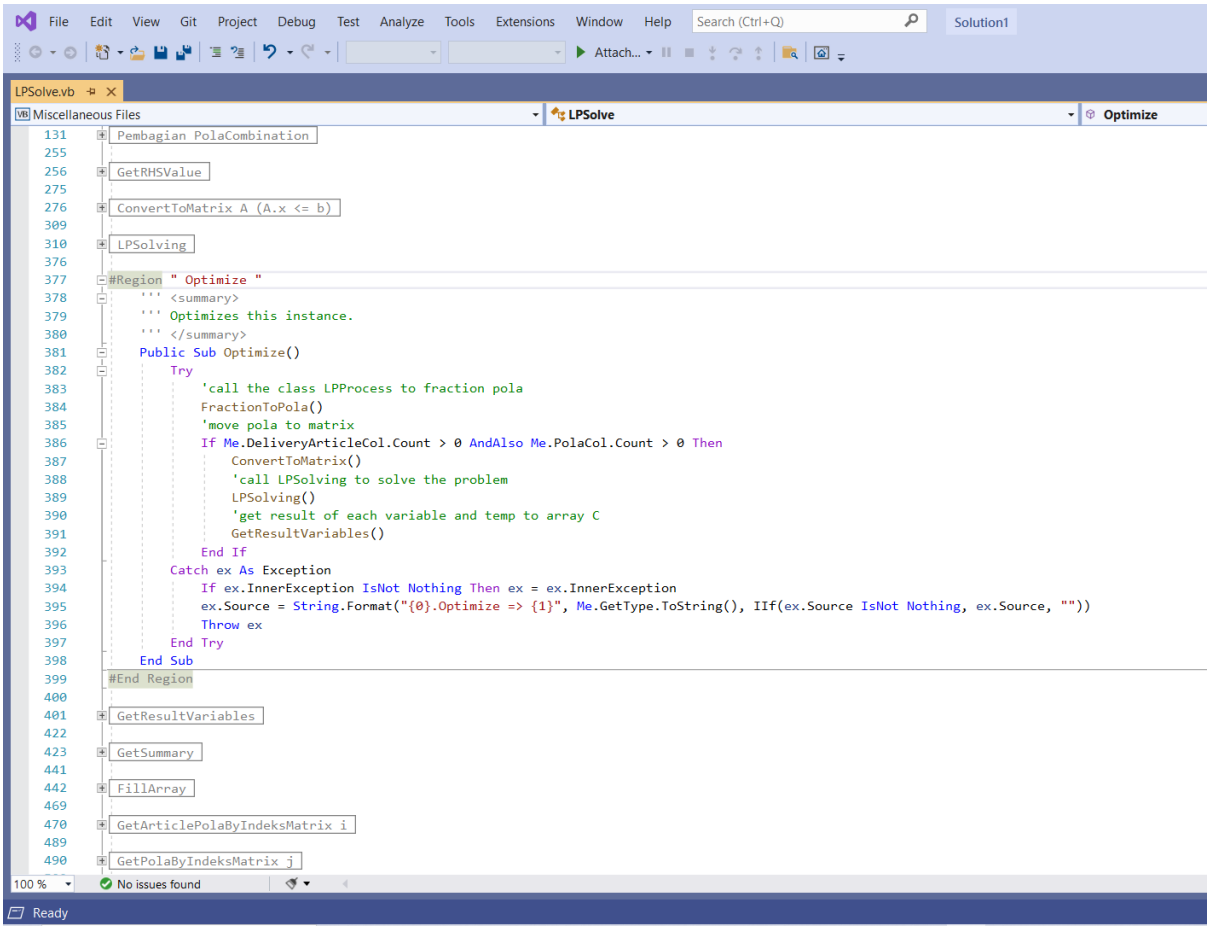
   Function ConvertToMatrix()



Function LPSolving()



Function Optimize()



⇨ I already have solved some study cases to hone and demonstrate the knowledge and skills that I have in data analytics and some in data science.

Please see the full Portfolio and other Portfolios at :
felicebenita.github.io/personal/portfolio.html

Berikut beberapa diantaranya :

2) Bellabeat Case Study

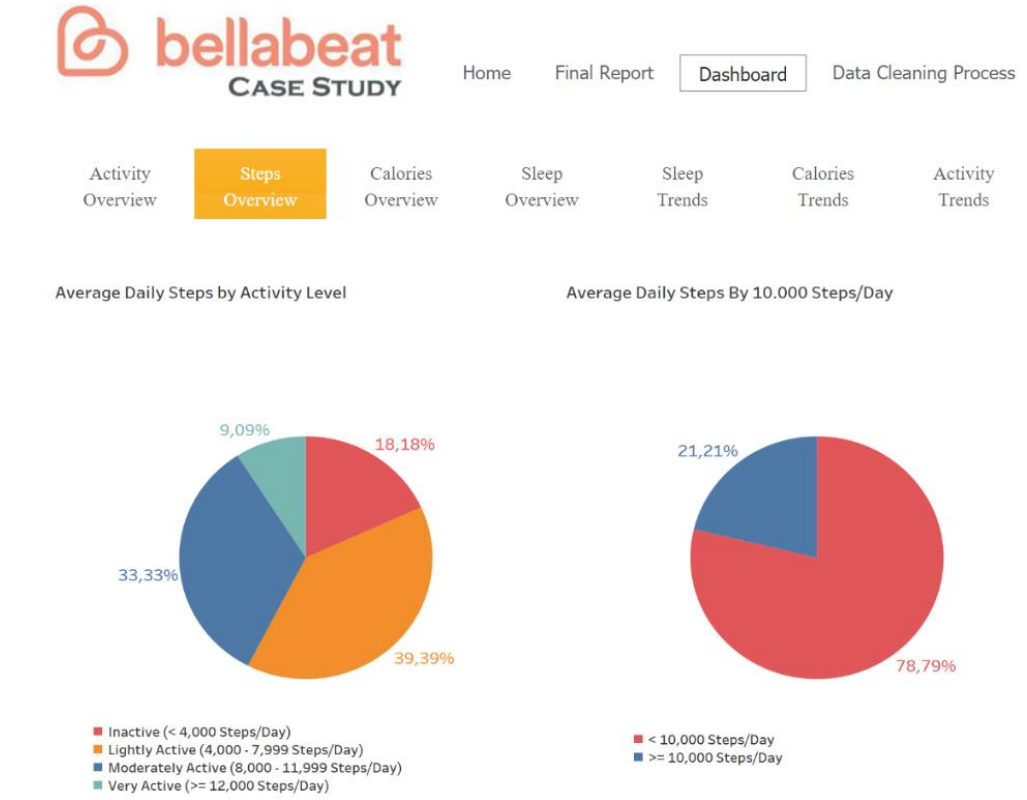
This case study was completed by me as part of the Google Data Analytics Professional Certificate capstone unit.

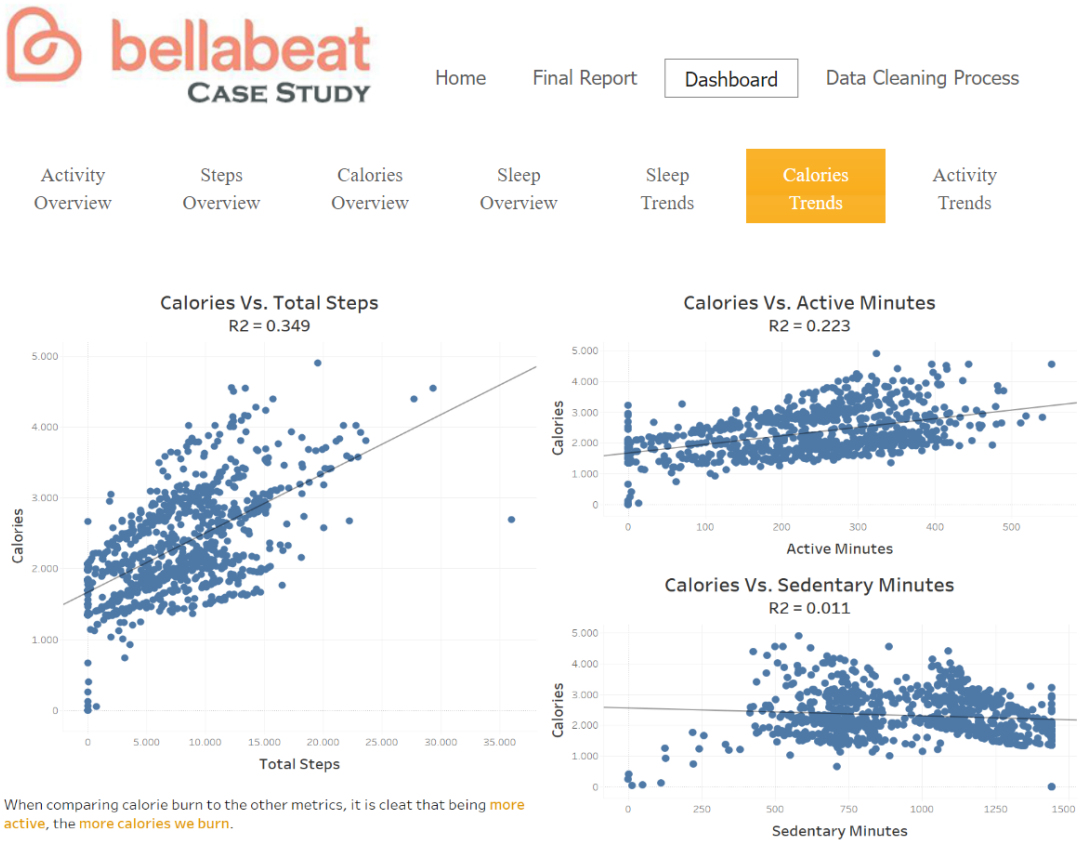
See Complete Portfolio : felicebenita.github.io/bellabeat-case-study

Summary

Business Task :
Analyze smart device fitness data in order to discover new growth opportunities for the company through improved marketing.

Some of the dashboards and reports :





6. Recommendations

Review: Bellabeat is a company that manufactures smart technologies for women to track varying aspects of their health. The company currently offers three different devices for tracking things such as activity, sleep, stress, and water intake. They also have an app (which users can use to see their health data) and a membership (which provides users with personalized health guidance).

The marketing recommendations for Bellabeat products are as follows :

1. Increase digital campaign showing the benefits to take at least 8,000 steps in a day for people's health

Average total steps per day are 7,638 which a little bit less for having health benefits for according to the CDC research. They found that taking 8,000 steps per day was associated with a 51% lower risk for all-cause mortality (or death from all causes). Taking 12,000 steps per day was associated with a 65% lower risk compared with taking 4,000 steps.

Bellabeat can encourage people to take at least 8,000 steps in a day by increasing digital campaign explaining the benefits for their health.

2. App reminders to remind users if there are factors that have not been achieved or excess in the recommended daily target

1) The recommended total steps in a day are at least 8,000 steps, if users still don't meet this target, the Bellabeat app can remind users to go for a run or walk.

Most activity happens between 5 pm and 7 pm - I suppose, that people go to a gym or for a walk after finishing work. Bellabeat can use this time to remind and motivate users to go for a run or walk.

2) If users consume daily calories less than or exceed their total daily calorie limit, the Bellabeat app can remind users to control their daily calorie consumption.

3) Adults are recommended to get between 7-9 hours of sleep a day on average as per the National Sleep Foundation's guidelines. Bellabeat should consider using app notifications to go to bed to ensure users get enough sleep, and sound an alarm when their sleep time exceeds the normal limit.

3. Provide a guide to the recommended daily calorie count and suggest some ideas for meals menu for those who want to lose weight or gain weight in order to achieve their target

If users want to lose weight, it's probably a good idea to control daily calorie consumption. Bellabeat can provide a recommended daily calorie guide and suggest some ideas for low-calorie meals menu.

It's the same with users who want to gain weight.

4. Create a feed in the app so members can share their short stories about achieving healthier life goals

With a kind of feed in the Bellabeat app that allows members to share short stories and their interesting experiences in achieving healthier life goals, this can motivate between members to improve their health levels and might help increase interest in the membership and increase retention.

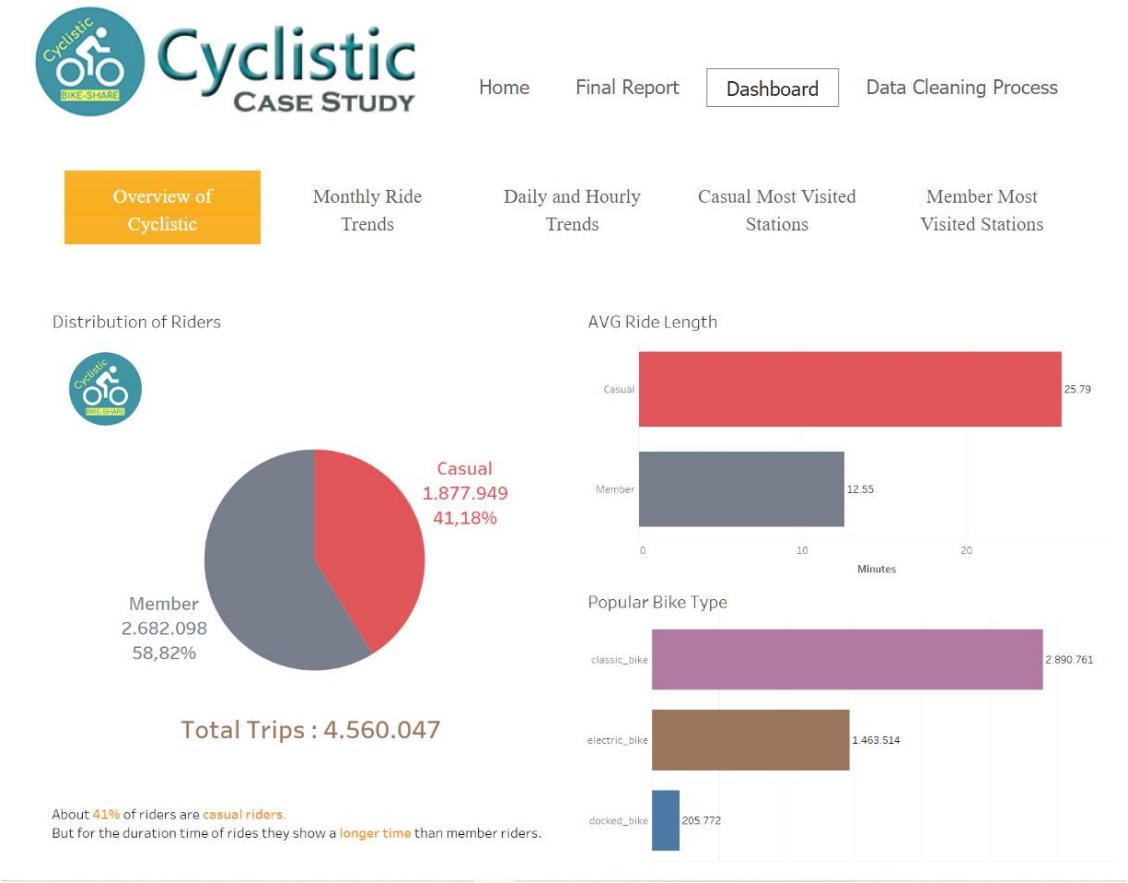
3) Cyclistic Case Study

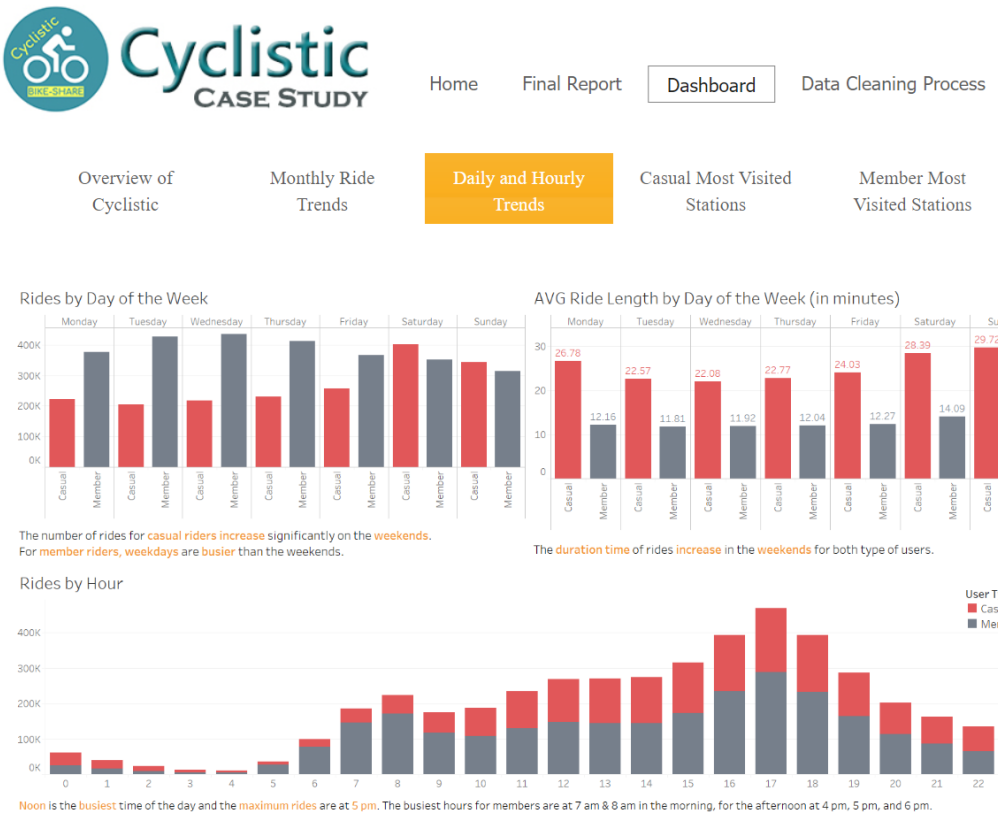
See Complete Portfolio : felicebenita.github.io/cyclistic-case-study

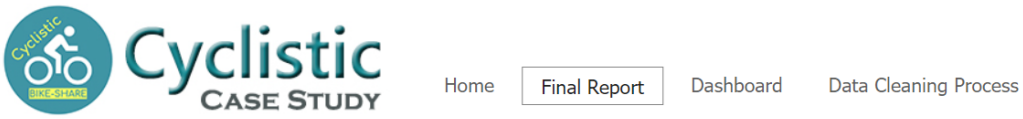
Summary

Business Task :
Analyze Cyclistic historical bike trip data to identify trends in order to design a new marketing strategy to convert casual riders into annual members.

Some of the dashboards and reports :







In summary, we have identified :

1. This indicates a different purpose in using Cyclistic services, where casual riders use bikes for leisure purposes while member riders use bikes for their work commutes.

The results of the analysis are shown in the graph as follows :

1) Most popular station

- The most popular station of casual riders shows that they usually ride bikes for commuting between recreational places (for leisure purpose).
- While the most popular station of member riders shows that they use bikes for commuting between work.

2) Most common day

- The number of rides for casual riders has increased significantly on the weekends, with the most common day is Saturday.
- For member riders weekdays are busier than the weekends, with the most common weekday is Wednesday.

3) Most popular time of day

- The busiest hours for members are at 7am & 8am in the morning, for the afternoon at 4pm, 5pm, and 6pm.

The significant increase of member riders using the bikes are at 7am & 8am and later on around 5pm indicates that member riders are using the Cyclistic bikes for their work commutes.

4) Average ride length

- The average ride length for casual riders is 25.79 minutes, while member riders for an average of 12.55 minutes.

The duration time of casual riders show a longer time than member riders. Again this further strengthens the argument that casual riders use Cyclistic bikes for leisure purposes.

2. Around 41% of riders are casual riders (not member). And the popular bike type is classic bike.

3. The busiest season is Summer and afternoon being the busiest time of the day.

6. Recommendations

As identified in the client brief, the marketing recommendations concluded from the insights of this case study should not be focused on encouraging new customers to use the Cyclistic bike service but instead focus on encouraging casual riders to convert to annual memberships.

The three possible marketing recommendations for Cyclistic are as follows :

1. Email reminders/notifications for casual riders to observe the price benefits and advantages of annual memberships

Offer the benefits and advantages that will be get if joining as an annual membership rather than regularly purchasing casual trips with Cyclistic. Email reminders and phone notifications should be used to remind regular casual riders of the long term pricing benefits from investing in an annual membership rather than purchasing regular casual trips.

2. Digital campaign which shows Cyclistic bike's being used in a Chicago local's everyday life

To encourage casual riders to become members, a digital campaign which encourages Chicago locals to observe how Cyclistic fits into their every day life would be beneficial. Also highlighting the benefits of riding bikes outdoors for health with good airflow. So that casual riders are interested in using bikes not only for leisure purposes.

3. Offer discounts for new membership and promos for the first 2 years membership

1) Offer annual membership price discounts (% off) for joining a new member.

New members will get a percentage discount (% off) from the regular annual membership price.

2) Offer attractive promos of special prices for the first 2 years of membership.

This will further grow loyalty in using the annual membership, because riders have been bound for the first 2 years using the annual membership. With this, the opportunity for riders to maintain their membership for the following years will be greater because they can feel the benefits after they join membership for a longer time.

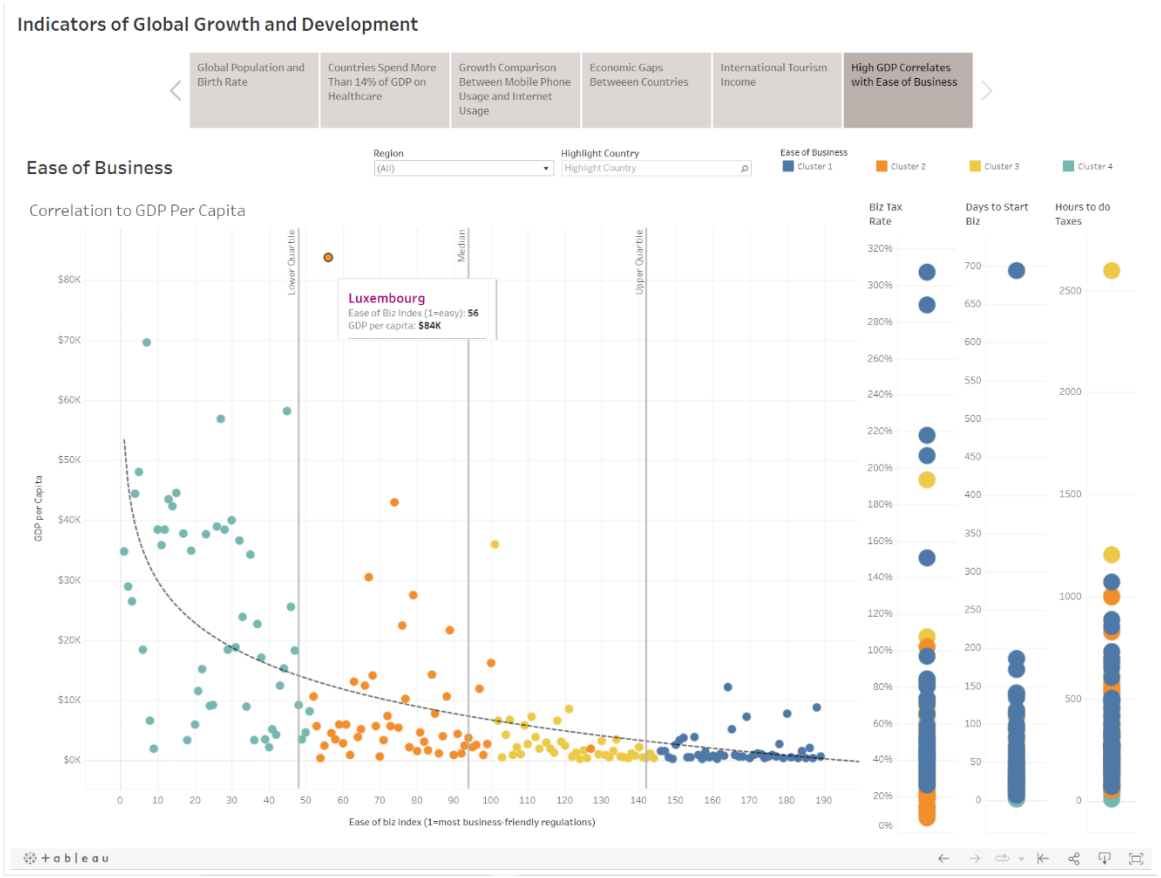
Email reminders should be used to offer non-members of this new membership discounts and promos.

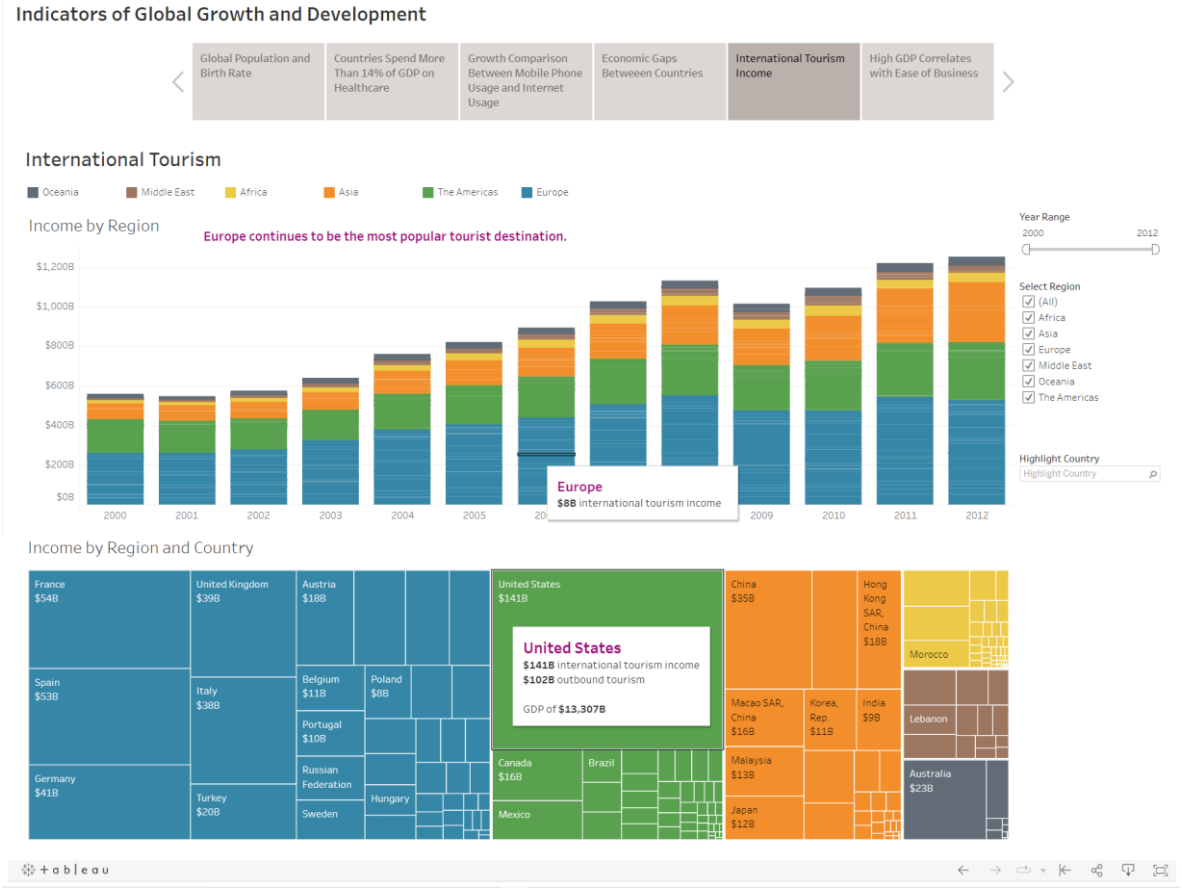
4) Indicators of Global Growth and Development

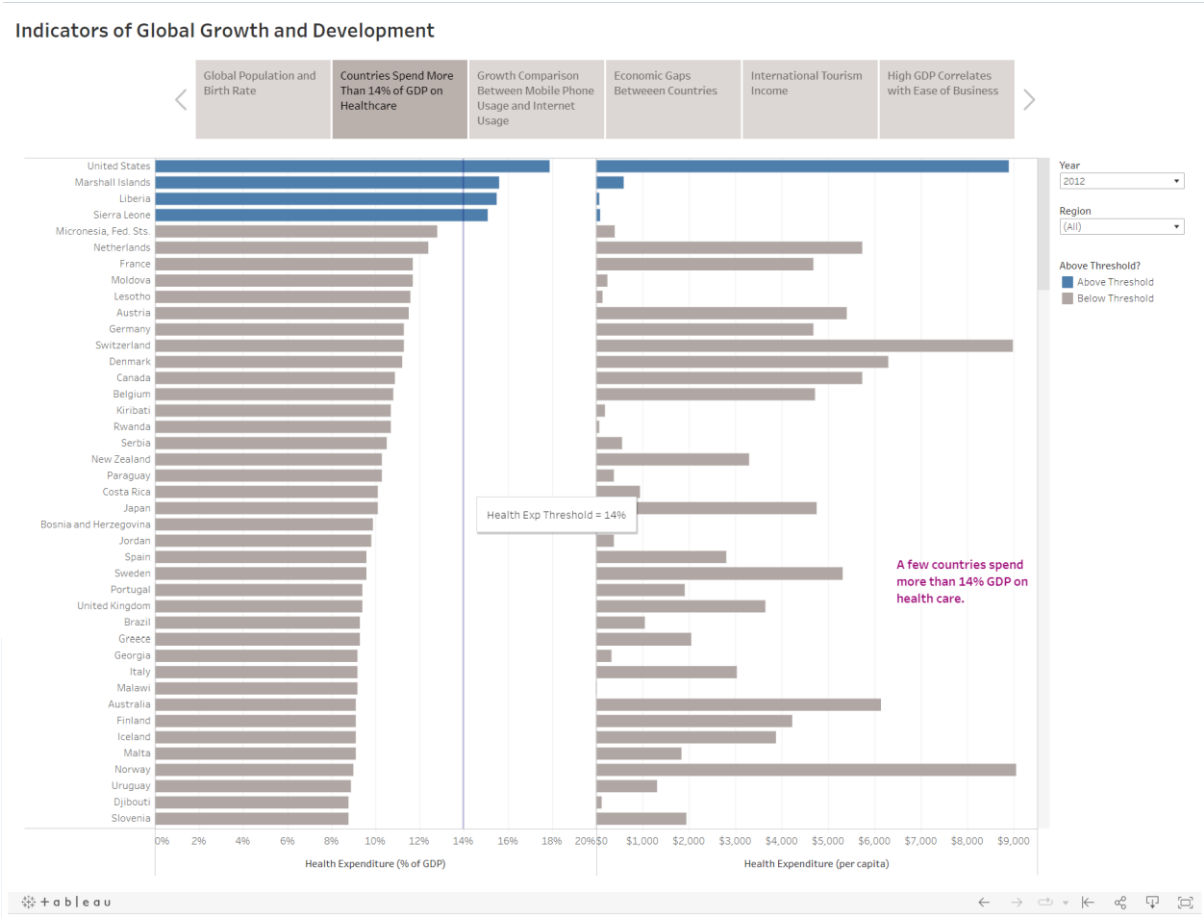
Build interactive dashboards and storytelling with Tableau to analyze the growth and development of countries in the world through several indicators, covering global population and birth rate; spends per Capita and percentage of GDP on healthcare; growth comparison between mobile phone usage and internet usage; comparison of GDP and GDP per Capita between countries; international tourism income by region and country; the correlation between high GDP per Capita with ease of business (and also show biz tax rate, days to start biz, and hours to do taxes).

See Complete Portfolio : felicebenita.github.io/personal/tableau/tableau-global-indicators.html

Some of the dashboards :





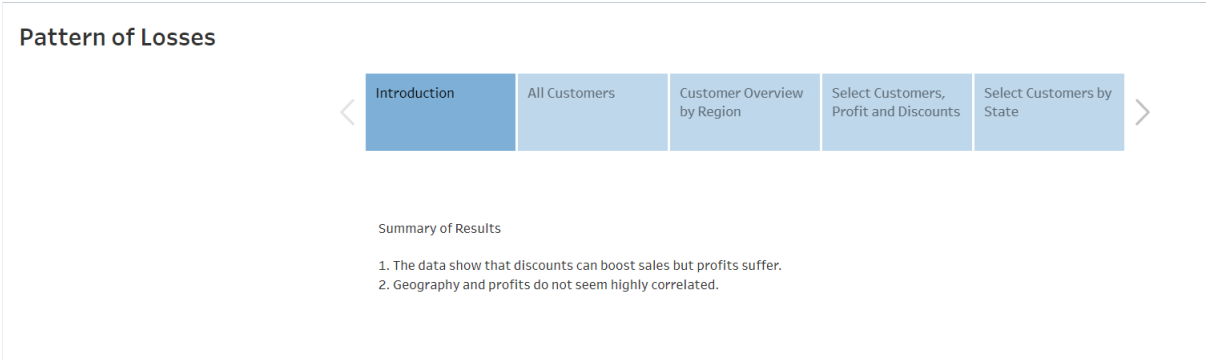


5) **Pattern of Losses**

Build interactive dashboards and storytelling with Tableau to analyze patterns of sales losses, what are the factors that affect sales losses, covering correlation between profit ratio, profit and sales by customer; correlation between customer profit and discounts; correlation between geography and profit.

See Complete Portfolio : felicebenita.github.io/personal/tableau/tableau-pattern-of-losses.html

Some of the dashboards :



Pattern of Losses

Introduction | All Customers | Customer Overview by Region | Select Customers, Profit and Discounts | Select Customers by State

==> Select on Bar of Each Region to Show Detail Customer Sales and Profit by Region

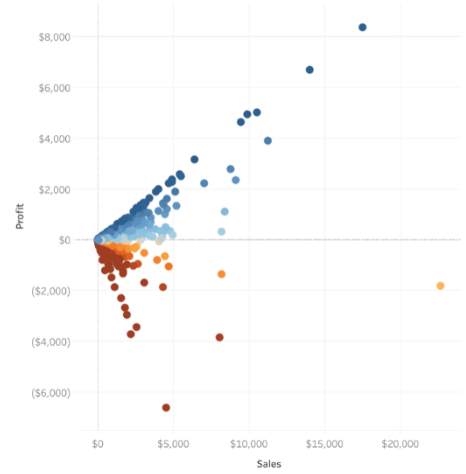
Customer Overview

Region	Sales per Customer	Sales	Quantity	Profit	Profit Ratio2
Central	$796.91	$501,256	8,780	$39,719	24158%
East	$1,007.17	$678,834	10,618	$91,521	47698%
South	$765.14	$391,750	6,209	$46,721	26523%
West	$1,057.60	$725,514	12,266	$108,386	70397%

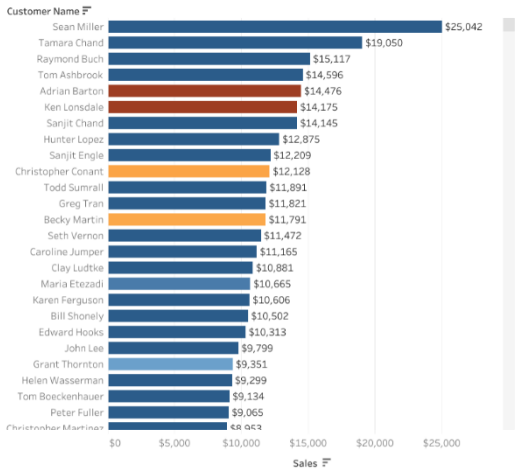
==> Hover on Circle of Scatter Plot to Highlights Customer Name on Custome Ranking

==> Select Customer Name to Show All Sales and Profit by Selected Customer on Scatter Plot

Sales and Profit by Customer



Customer Ranking



Year: (All)

Category: (All), Furniture, Office Supplies, Technology

Segment: (All), Consumer, Corporate, Home Office

Profit Ratio: -50% 50%

Pattern of Losses

Introduction | All Customers | Customer Overview by Region | Select Customers, Profit and Discounts | Select Customers by State

==> Select on Bar of Each Customer on Customer Discounts to Show Selected Customer Profit Detail

Customer Profit

Category	Sub-Category	Aaron Smayling	Sharelle Roach	Grant Thornton	Raymond Buch	Tom Ashbrook	Aaron Bergman	Aaron Hawkins	Adam Bellavance	Adam Hart	Adam Shillingsb..	Adrian Barton	Adrian Ha
Furniture	Bookcases	84	57				55				-44		
	Chairs		-1				5	21	-73	-35	84		-26
	Furnishings				25	-63		38	41	137	37	130	
	Tables	-6		-347					367	-142	215	-204	-264
Office Supplies	Appliances			15	50				45	1	4		
	Art		33		98	2	5	2		4	16	5	3
	Binders	-93	-32	59	-9	149		22	1,437	289	-337	5,389	10
	Envelopes					29		121			25		
	Fasteners									9			
	Labels				10			9		2	4	2	5
	Paper	45	7		10			4	166	55	8		27

-10K 0K 10K — Profit

Customer Discounts

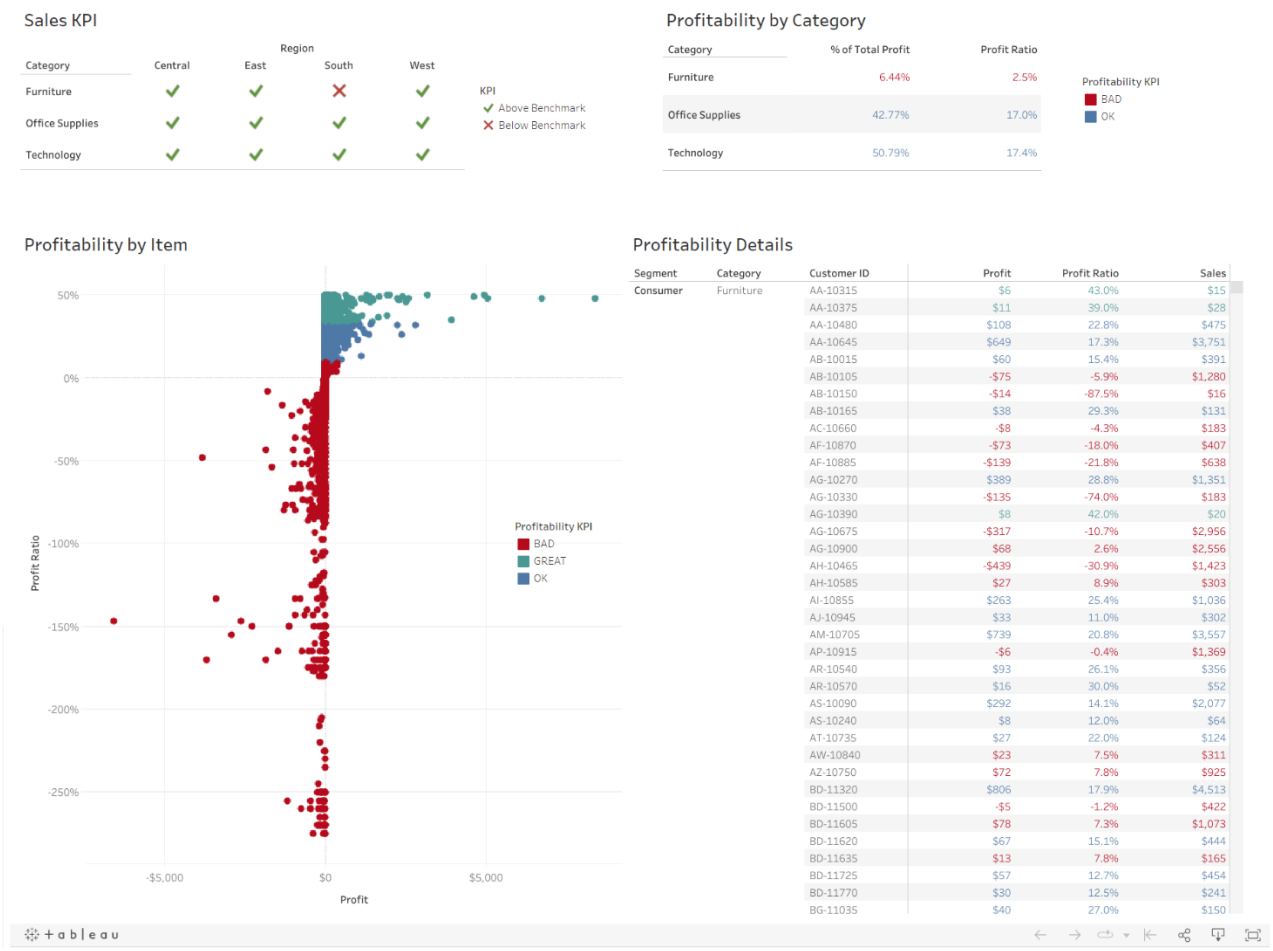
Category	Sub-Category	Tony Sayre	Toby Swindell	Toby Gnade	Michelle Moray	Lisa Hazard	Justin Ritter	Juliana Krohn	Jim Radford	Jeremy Farry	Irene Maddox	Corey Catlett	Bill Stewart
Furniture	Chairs	20%		20%	20%			0%		17%	0%	20%	20%
	Bookcases	70%	70%	70%	70%	70%	70%	70%	70%	70%	70%	70%	70%
	Tables				50%						50%	0%	
	Furnishings	10%	0%			0%	20%	0%			20%	0%	20%
Office Supplies	Binders	20%	38%	20%	45%	75%	35%	35%		20%	30%	70%	45%
	Paper	20%	10%	20%	0%		0%			7%	13%		
	Art				10%		20%	20%	20%	0%	20%	20%	20%
	Supplies							20%			0%		
	Storage	20%	0%	20%	10%				0%	20%	0%	20%	10%
	Labels			20%	0%					0%	0%		
	Fasteners									10%	0%		20%
	Envelopes		20%			20%							
	Appliances	20%			0%							20%	
Technology	Machines												
	Phones			0%	20%	40%	20%	0%			0%	10%	20%
	Copiers			0%									40%
	Accessories	0%			0%		7%			20%	20%		20%

6) **Sales KPI**

Sales KPI (Key Performance Indicator) visualization in Tableau that shows sales by category and item.

See Complete Portfolio : felicebenita.github.io/personal/tableau/tableau-sales-kpi.html

Some of the dashboards :



7) Netflix Movies and TV Shows

Exploratory Data Analysis with Visualization in Python using Netflix Movies and TV Shows data.

See Complete Portfolio : felicebenita.github.io/personal/jupyter-notebook/netflix-eda-visualization.html

Some of codes & visualizations :

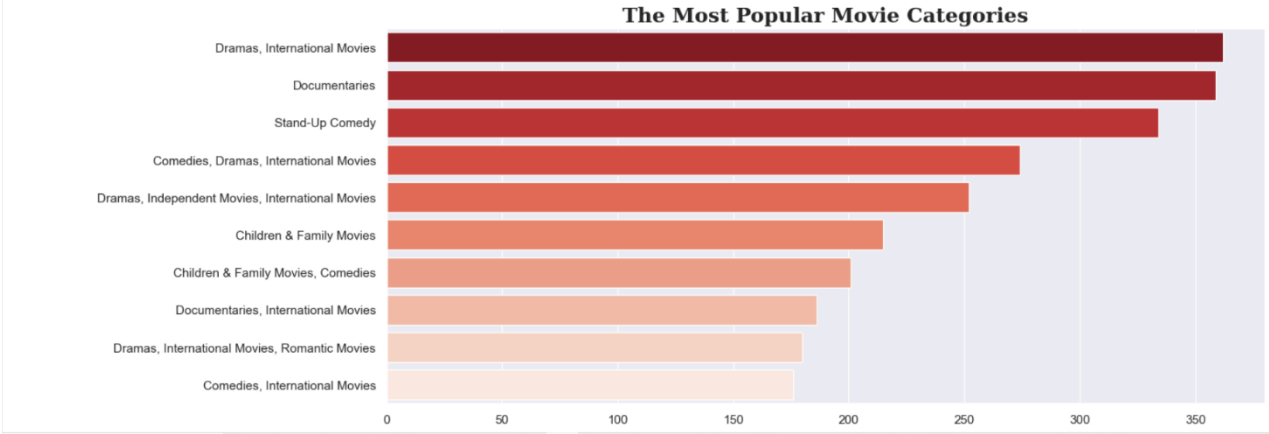
4.2. The Most Popular Categories

4.2.1 The Most Popular Movie Categories

```
In [12]: movies_cat = df[df['type'] == 'Movie']
movies_cat = movies_cat[movies_cat['listed_in'] != 'Unspecified']
movies_cat = movies_cat.groupby('listed_in').agg({'show_id':'count'}).rename({'show_id': 'total_cat'}, axis = 1)
movies_cat = movies_cat.sort_values(['total_cat'], ascending = (False)).head(10).reset_index()

sns.set(rc = {'figure.figsize': (14, 6)})
ax = sns.barplot(data=movies_cat, x="total_cat", y="listed_in", palette="Reds_r")
ax.set_title("The Most Popular Movie Categories", fontfamily='serif', fontsize=18, fontweight='bold')
ax.set(xlabel=None, ylabel=None)

Out[12]: [Text(0.5, 0, ''), Text(0, 0.5, '')]
```

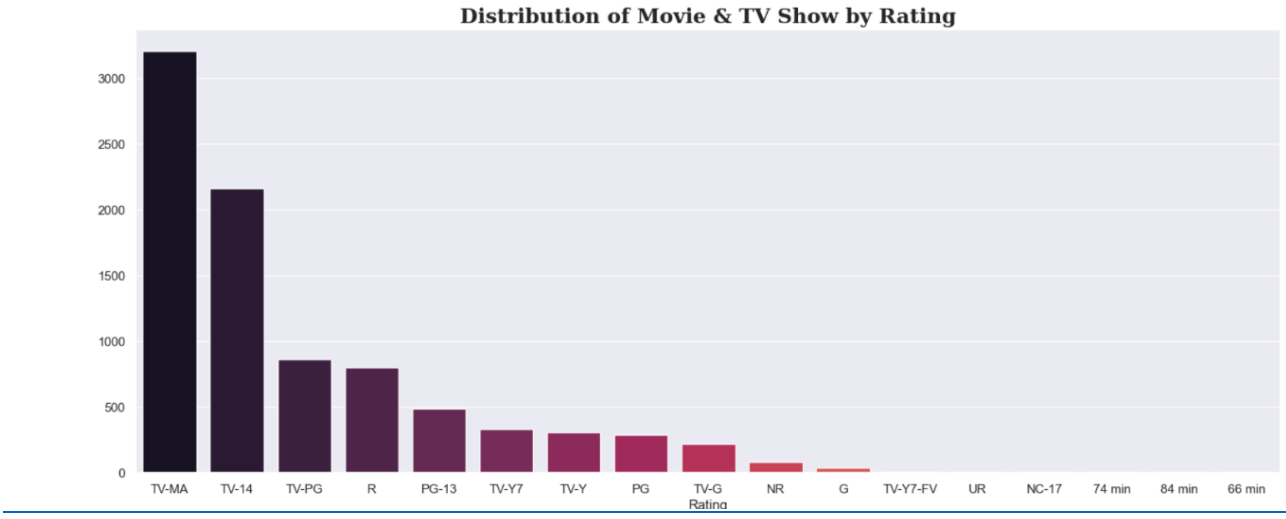


4.8. Distribution of Movie & TV Show by Rating

```
In [20]: ratings = df.groupby(['rating']).agg({'show_id':'count'}).rename({'show_id': 'total_shows'}, axis = 1)
ratings = ratings.sort_values(['total_shows'], ascending = (False)).head(20).reset_index()

sns.set(rc = {'figure.figsize': (18, 7)})
ax = sns.barplot(data=ratings, x="rating", y="total_shows", palette='rocket')
ax.set_title("Distribution of Movie & TV Show by Rating", fontfamily='serif', fontsize=18, fontweight='bold')
ax.set(xlabel='Rating', ylabel=None)

Out[20]: [Text(0.5, 0, 'Rating'), Text(0, 0.5, '')]
```



8) **Goodreads Books**

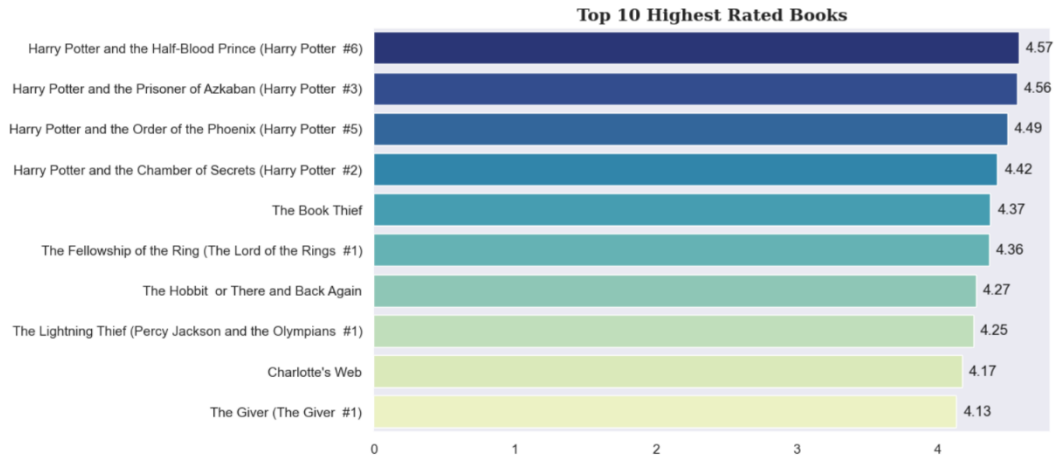
Exploratory Data Analysis with Visualization in Python using Goodreads Books data.

See Complete Portfolio : felicebenita.github.io/personal/jupyter-notebook/goodreads-eda-visualization.html

Some of codes & visualizations :

4.1. Top 10 Highest Rated Books

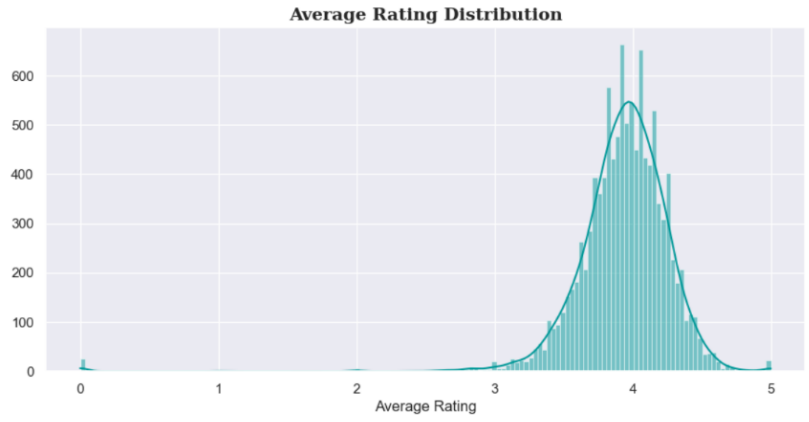
In [12]:
```
# Visualize Top 10 Highest Rated Books
top_books = top_books.sort_values(by='average_rating', ascending=False).head(10)
sns.set(rc = {'figure.figsize': (10, 6)})
ax = sns.barplot(data=top_books, x="average_rating", y="title", palette="YlGnBu_r")
ax.set_title("Top 10 Highest Rated Books", fontfamily='serif', fontsize=14, fontweight='bold')
ax.set(xlabel=None, ylabel=None)
for i in ax.patches:
    ax.text(i.get_width() + .05, i.get_y() + 0.5, str(i.get_width()), fontsize = 12, color = 'k')
ax.grid(False)
```



4.9. Average Rating Distribution

In [22]:
```
# Average Rating Distribution
df.average_rating = df.average_rating.astype(float)
sns.set(rc = {'figure.figsize': (11, 5)})
ax = sns.histplot(data=df['average_rating'], kde=True, color='#099')
ax.set_title("Average Rating Distribution", fontfamily='serif', fontsize=14, fontweight='bold')
ax.set(xlabel='Average Rating', ylabel=None)
```

Out[22]:
```
[Text(0.5, 0, 'Average Rating'), Text(0, 0.5, '')]
```



Some of Data Science Study Case :

1) **Human Activity Recognition Using Smartphones Data**
   Context : We will be using the Human Activity Recognition with Smartphones database, which was built from the recordings of study participants who carried a smartphone with an embedded inertial sensor while performing activities of daily living (ADL). The objective is to classify the activities the participants performed into one of the six following categories: walking, walking upstairs, walking downstairs, sitting, standing, and laying.
   Covers all the necessary phases for the machine learning algorithms used i.e. retrieving data, data wrangling, EDA and feature engineering, model development, and classification error metrics.

   See Complete Portfolio : felicebenita.github.io/personal/jupyter-notebook/human-activity-recognition.html

2) **Credit Risk Prediction**
   Context : We will be using the German Credit Data, which each entry represents a person who takes a credit by a bank. The objective is to classified each person as good or bad credit risks according to the set of attributes.
   Covers all the necessary phases for the machine learning algorithms used i.e. retrieving data, data wrangling, exploratory data analysis, feature engineering and variable transformation, model development, dan classification error metrics.

   See Complete Portfolio : felicebenita.github.io/personal/jupyter-notebook/credit-risk-prediction.html

3) **Customer Segmentation**
   Context : An automobile company has plans to enter new markets with their existing products (P1, P2, P3, P4 and P5). After intensive market research, they've deduced that the behaviour of new market is similar to their existing market. In their existing market, the sales team has classified all customers into 4 segments (A, B, C, D). The objective is to classified new potential customers into 4 segments similar to their existing market.
   Using K-Means clustering algorithms, which one of the most customer clustering algorithms. It relies on finding cluster centers to group data points based on minimizing the sum of squared errors between each datapoint and its cluster center.

   See Complete Portfolio : felicebenita.github.io/personal/jupyter-notebook/customer-segmentation.html