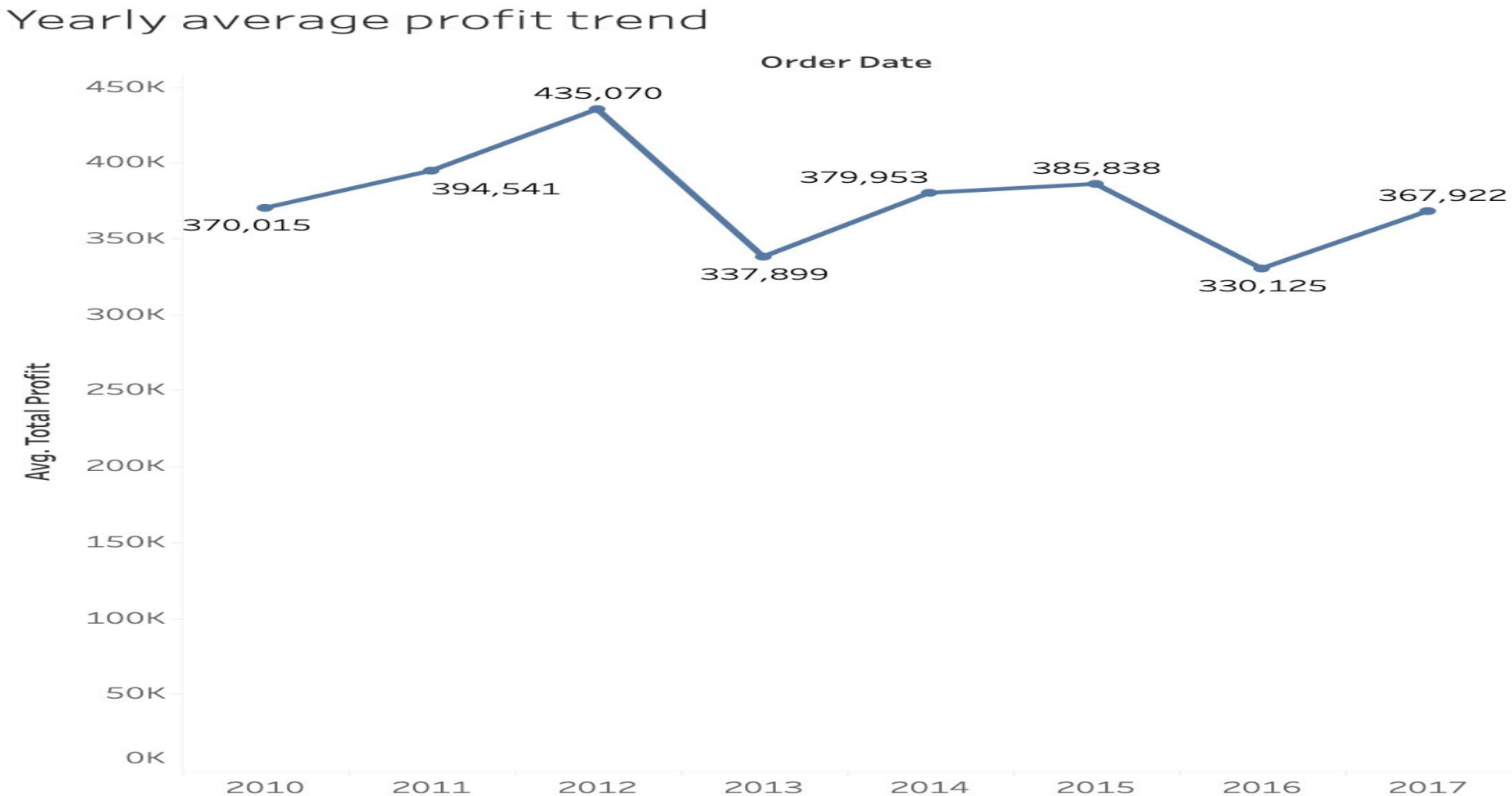
--====Europe sales analysis using SQL, Tableau ,Python and Excel along with Insights and recommendations====--

Bala Mira G

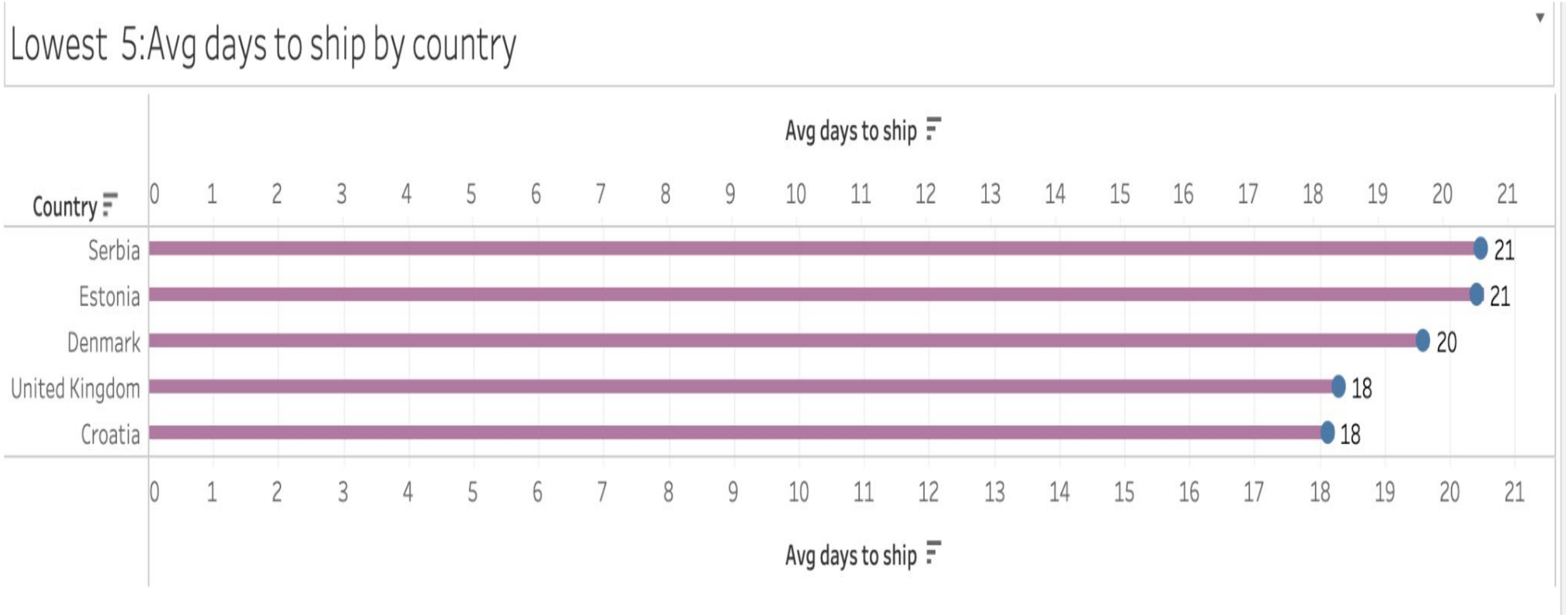
----- EDA Insights====--

--=====Yearly Avg Profit =====-

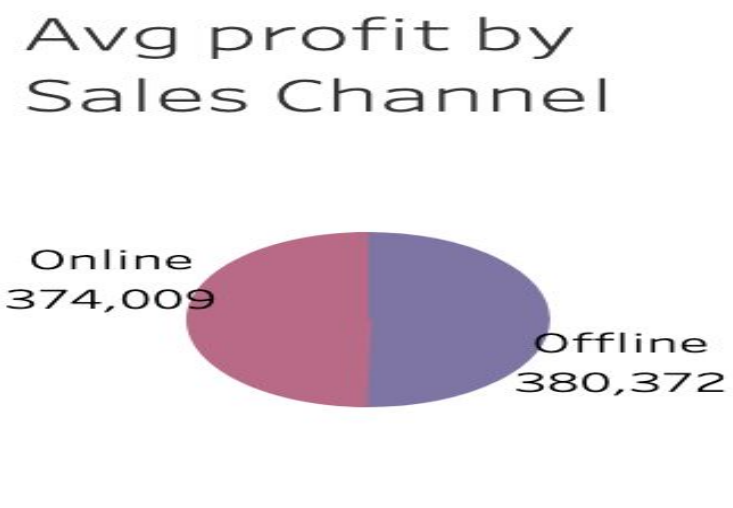


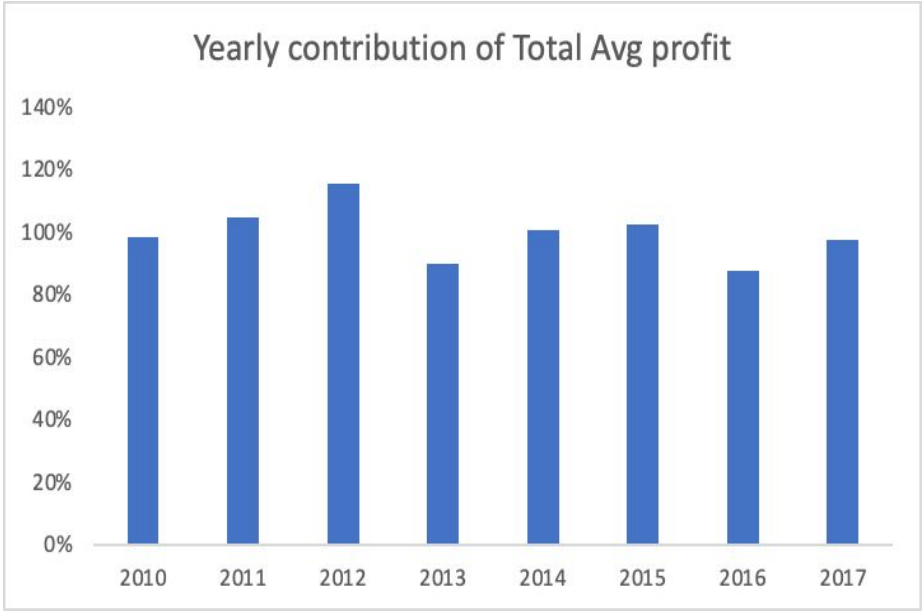
--====EDA Insights===--

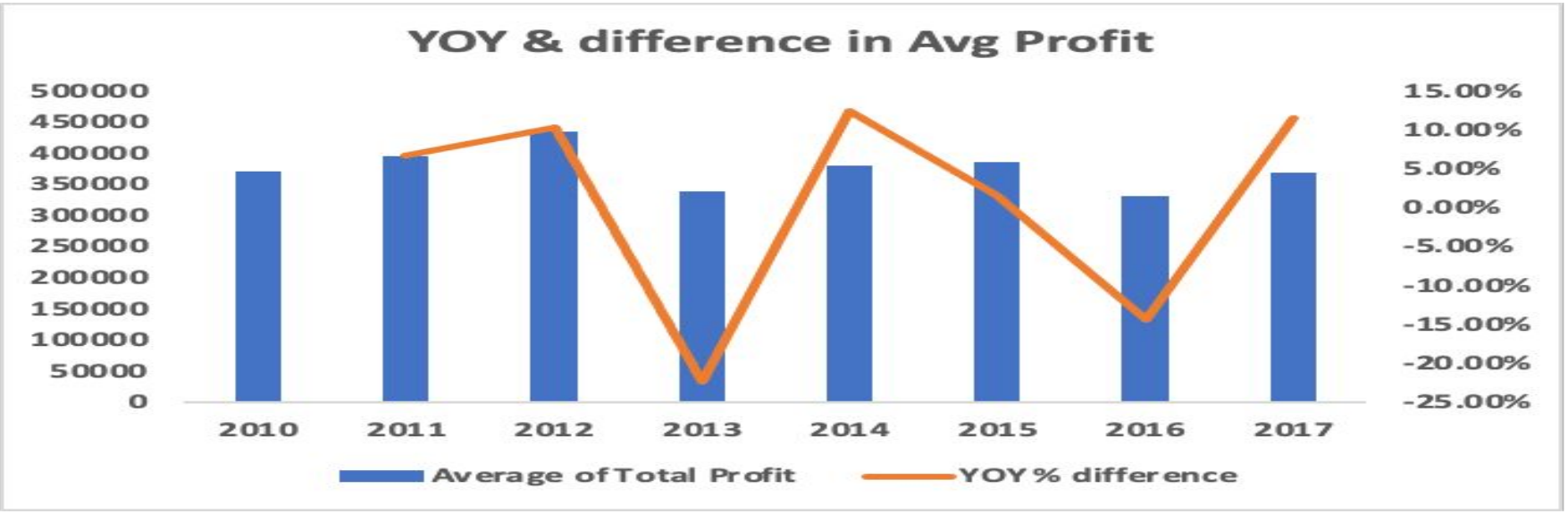
-------- 5 countries with lowest Avg time to Ship =====-



--===EDA Insights===--

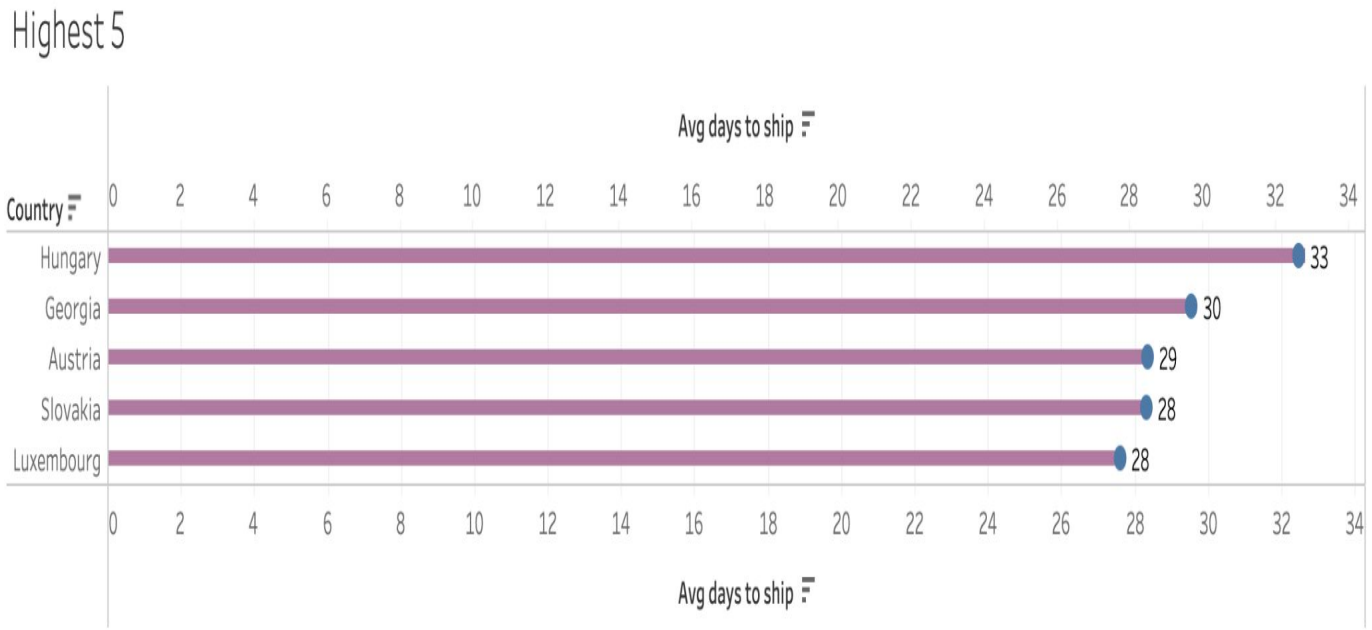


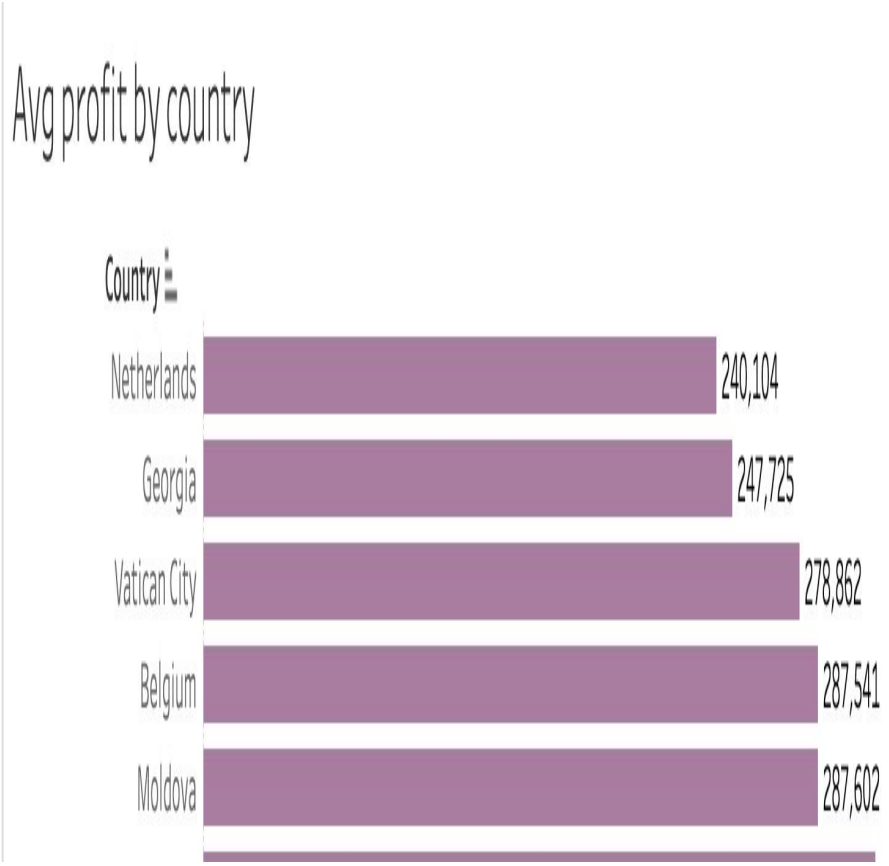




--====EDA Insights===--

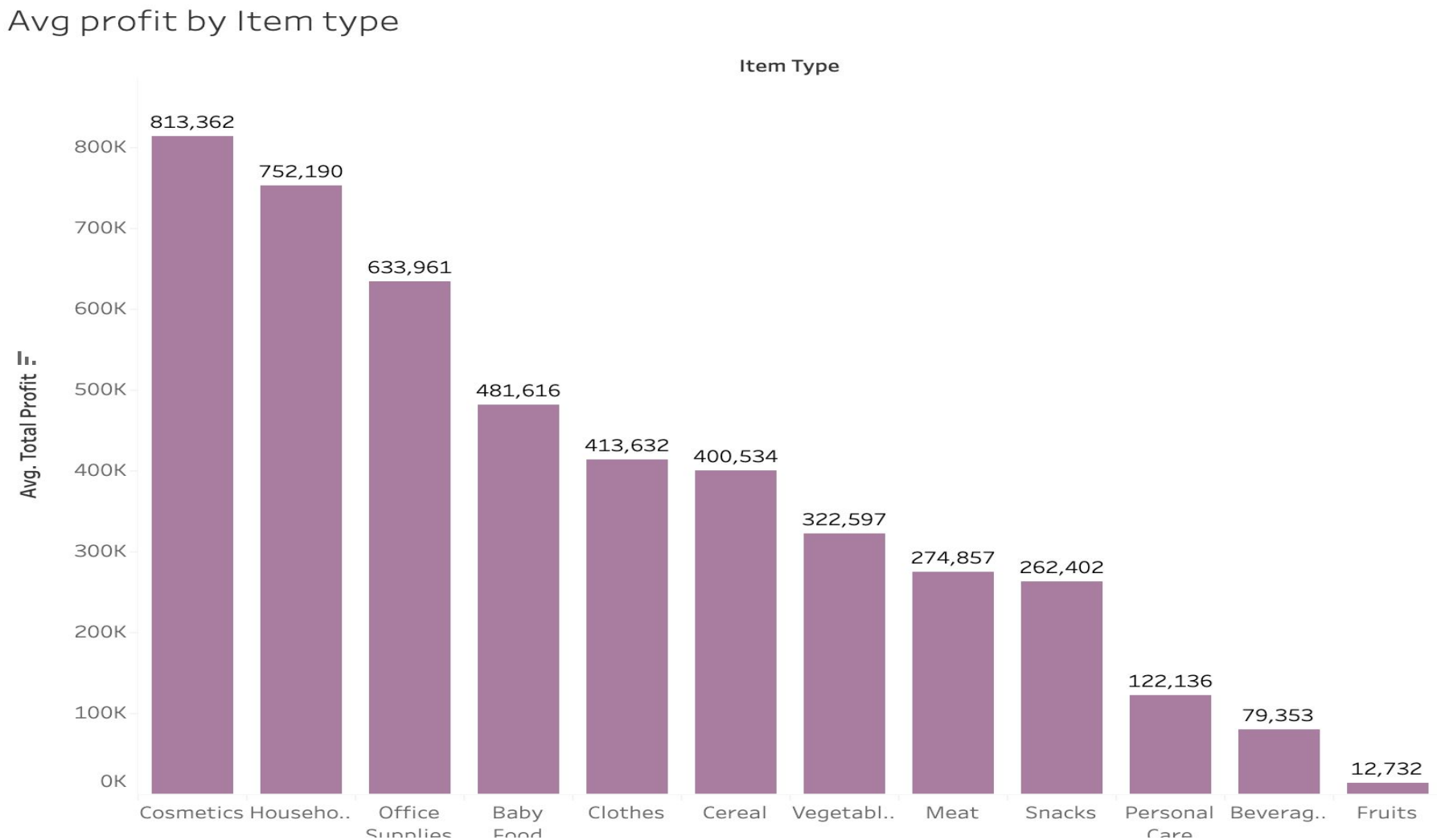
------ 5 countries with highest Avg time to Ship and lowest Avg profit by country =====-





-----EDA Insights====--

--=====Avg profit by Item type =====-

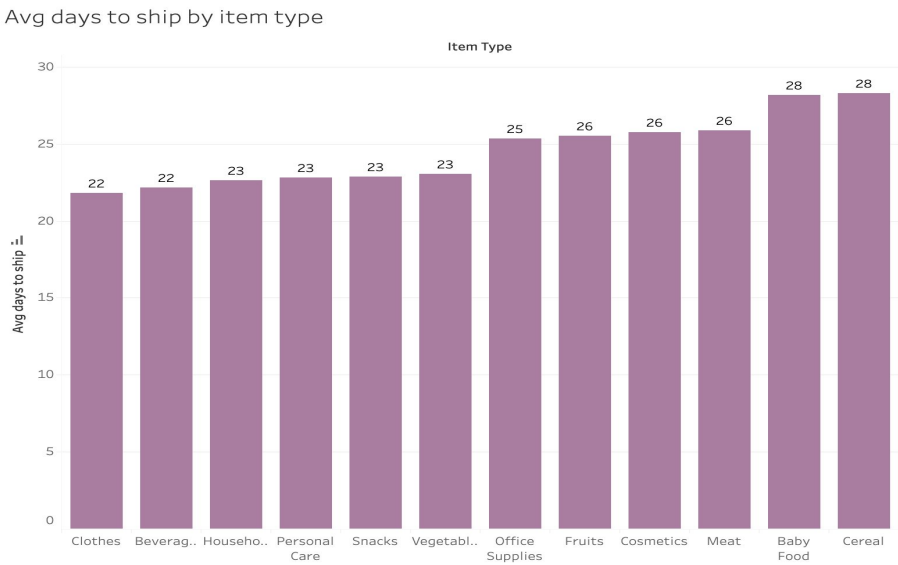


-----Insights===--

--===== EDA insights =====-

```sql
#Avg profit by item type#

with cte as(SELECT
    ROUND(AVG(`Total Profit`), 0) AS avgprofit, ROUND(AVG(`Total Cost`), 0) AS avgcost, `Item type`
FROM
    europe
GROUP BY `Item type`)
select * from cte where avgprofit>avgcost;
```

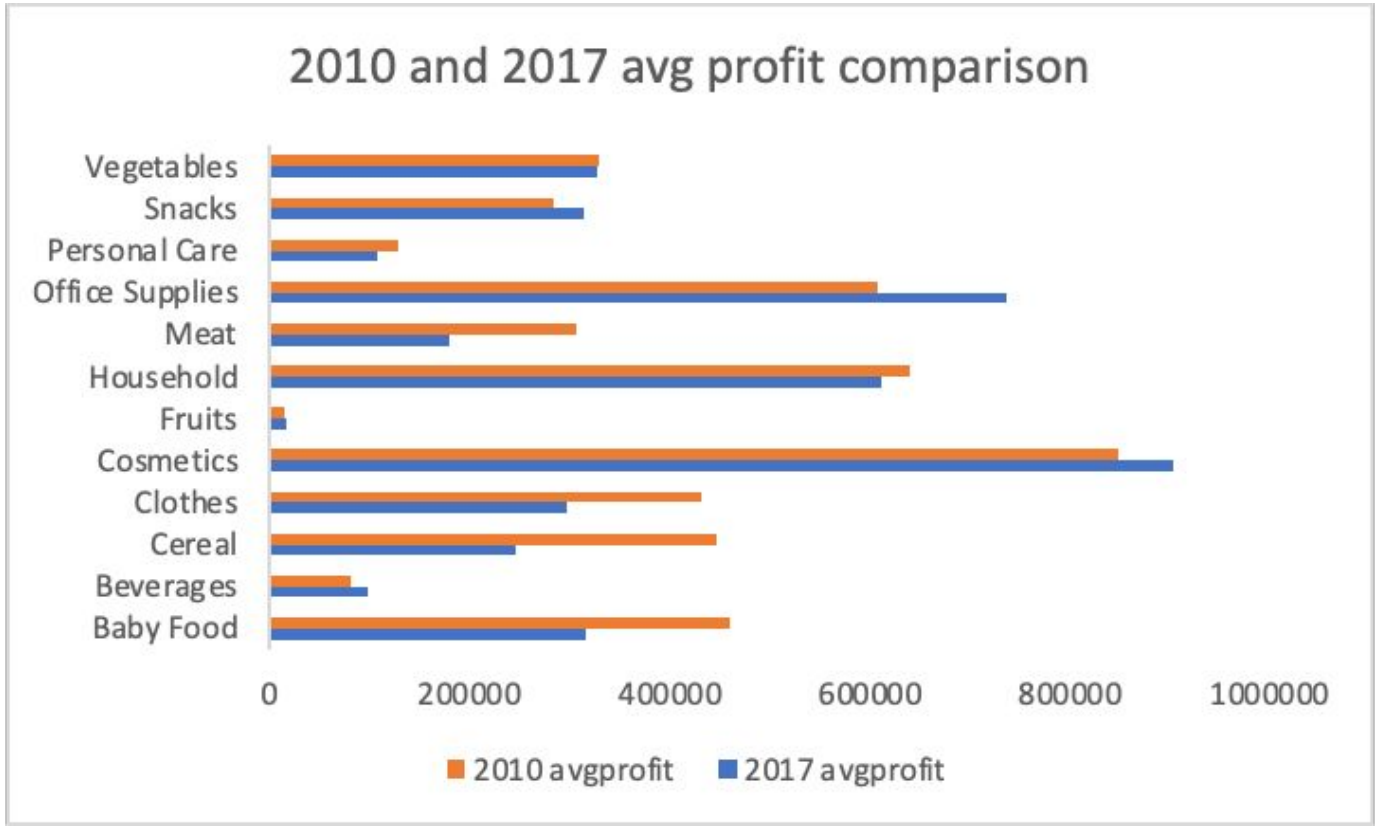
avgprofit	avgcost	Item type
413632	201859	Clothes



Insights:

-

--===EDA Insights===--



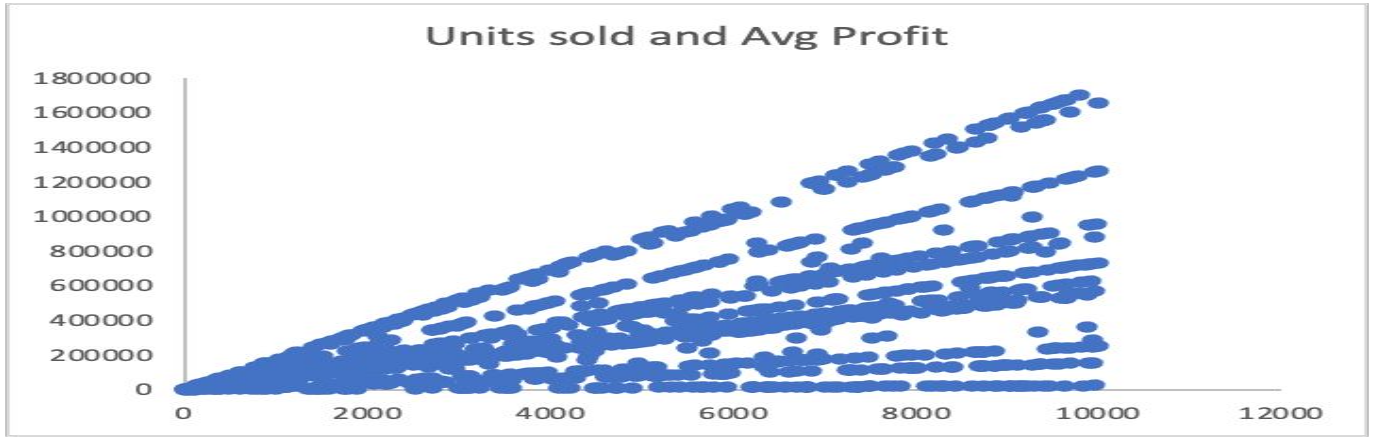
Insights:

- Beverages, Cosmetics ,Fruits, Office, Supplies and Snacks are those whose profit in 2010 is lower than that of 2017.
- Steps can be taken for the other Item types to increase their profit in future.

-----Insights=====--

------ Correlation between Avg profit and Units sold====-

	Units Sold	Avg Profit
Units Sold	1	
Avg profit	0.59474029	1



<u>Business case scenario</u>:

The business wants to introduce a “Combo sale season” and would want to check if increasing the number of units might increase the average profit.

- We did a correlation analysis between these two variables to analyze the historical data.
- The correlation coefficient between units sold and average profit is approximately 0.5947.
- This value indicates a moderate positive relationship. As the number of units sold increases, the average profit also tends to increase.
- However, correlation does not imply always causation. Though there is a positive correlation, it does not necessarily mean that increasing the units sold causes an increase in average profit. We can further analyze this by using Regression analysis.The output and interpretation of the results are mentioned in the following slides.

----Insights===--

--==== Regression analysis between Avg profit and Units sold=====-

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.59474028
R Square	0.353716
Adjusted R Square	0.35319396
Standard Error	294736.215
Observations	1240

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	5.886E+13	5.886E+13	677.566534	1.708E-119
Residual	1238	1.0754E+14	8.6869E+10		
Total	1239	1.664E+14			

.

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	8845.0427	16539.5852	0.5347802	0.59289789	-23603.673	41293.7579	-23603.673	41293.7579
X Variable 1	74.8663941	2.8761461	26.0301082	1.708E-119	69.2237347	80.5090535	69.2237347	80.5090535

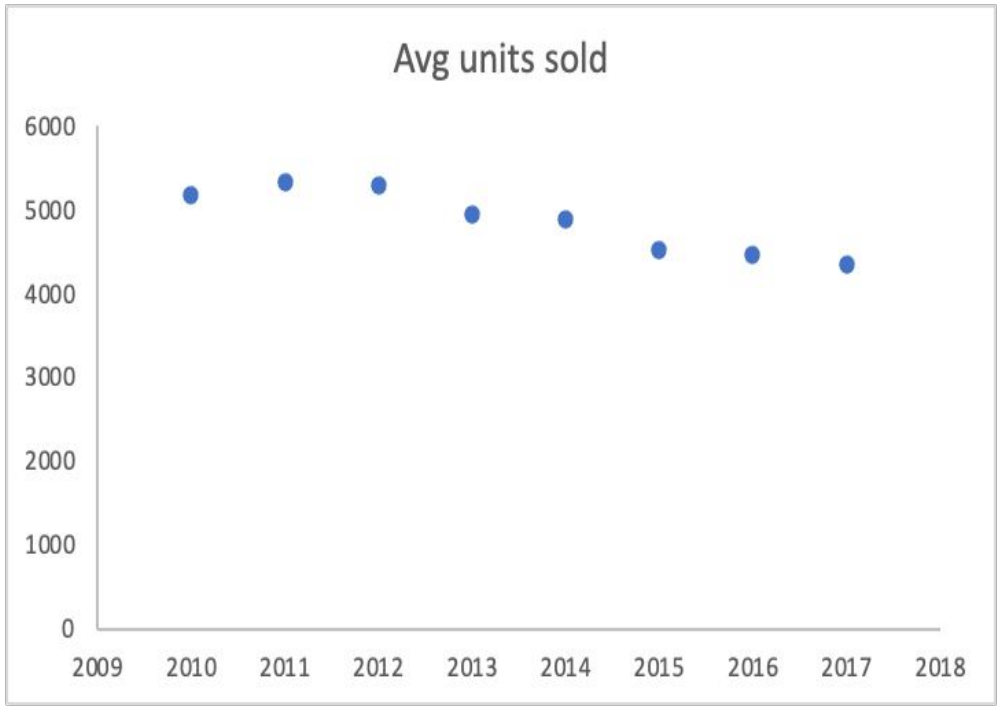
- **Multiple R**: This is the correlation coefficient between the independent and dependent variables. In this case, it is 0.5947, suggesting a moderate positive correlation.
- **R Square** (Coefficient of Determination): Approximately 35.37% of the variance in the dependent variable (average profit obtained) can be explained by the independent variable (avg units sold). This indicates a moderate fit of the model.
- **Correlation:** There's a moderate positive connection between the number of units sold and the profit.
- **Explanation:** About 35% of the changes in profit can be explained by the changes in the number of units sold.
- **Model Significance:** The overall model (relationship) is quite strong and not likely due to chance.
- **Prediction:** For each additional unit sold, the profit is expected to increase by around 75 units.

-----Insights=====--

------ Regression to plan inventory stocks=====-

Business case scenario:

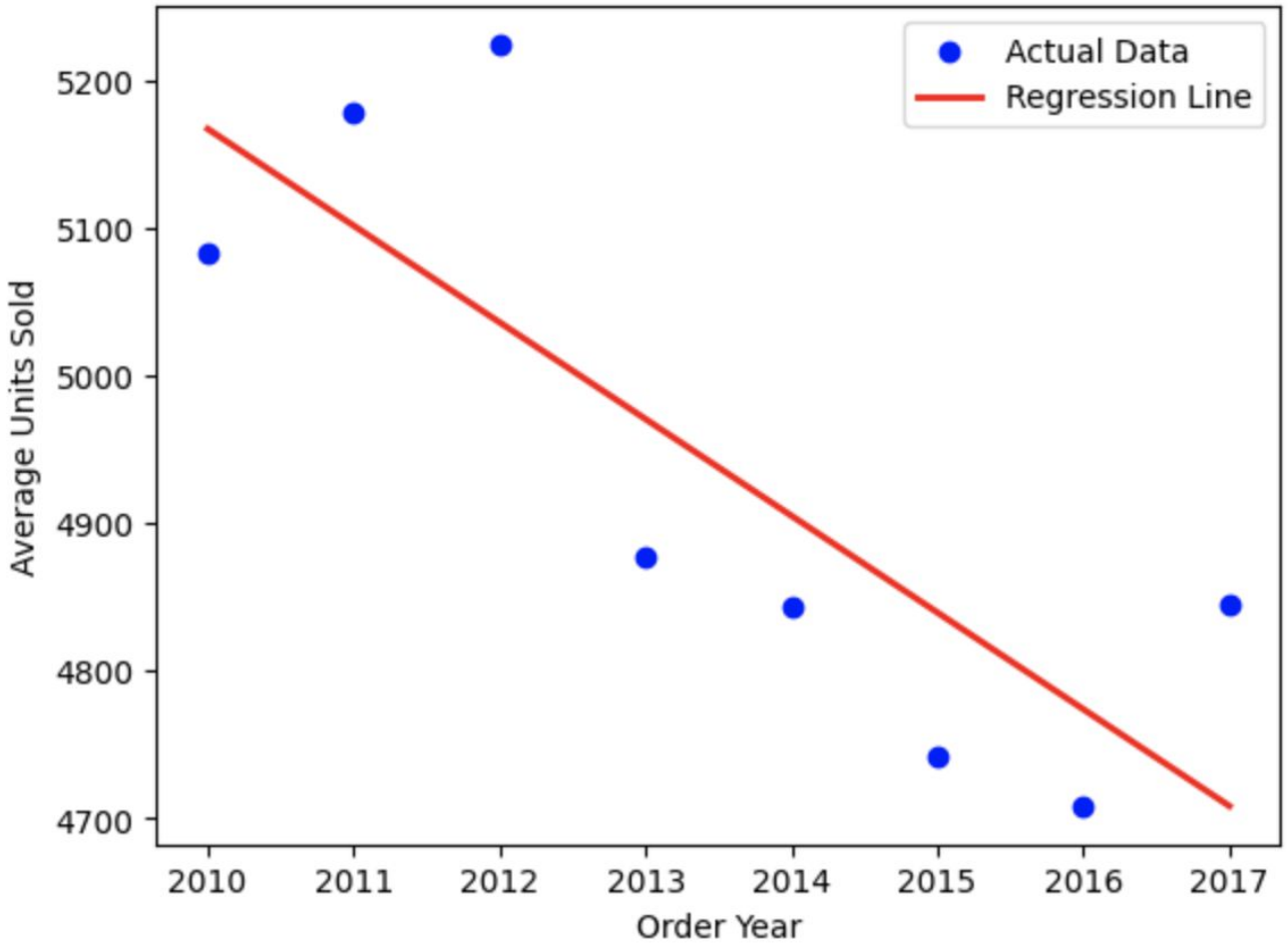
The business wants to plan the inventory levels by predicting if units sold would go up or go down. We could use regression analysis in either Python or Excel to get the desired results.




```python
import pandas as pd
import statsmodels.api as sm

# Provided data
data = {
    'Year': [2010, 2011, 2012, 2013, 2014, 2015, 2016, 2017],
    'Inventory': [5184, 5339, 5283, 4948, 4881, 4525, 4458, 4340],
}

df = pd.DataFrame(data)

# Define independent variable (X) and dependent variable (y)
X = sm.add_constant(df['Year'])
y = df['Inventory']

# Fit the regression model
model = sm.OLS(y, X).fit()

# Print the regression results
print(model.summary())
```

-----Insights=====--

------ Regression to plan inventory stocks=====-



- **OLS Regression Results :**
- ==========
- Dep. Variable: Avg_Units_Sold R-squared: 0.660
- Model: OLS Adj. R-squared: 0.603
- Method: Least Squares F-statistic: 11.62
- Date: Sat, 02 Mar 2024 Prob (F-statistic): 0.0143
- Time: 22:16:29 Log-Likelihood: -48.798
- No. Observations: 8 AIC: 101.6
- Df Residuals: 6 BIC: 101.8
- Df Model: 1
- Covariance Type: nonrobust
- ==========
- coef std err t P>|t| [0.025 0.975]
- ---
- const 1.368e+05 3.87e+04 3.537 0.012 4.22e+04 2.32e+05
- Order_Year -65.5119 19.217 -3.409 0.014 -112.534 -18.490
- ==========
- Omnibus: 2.059 Durbin-Watson: 1.733
- Prob(Omnibus): 0.357 Jarque-Bera (JB): 1.147
- Skew: 0.680 Prob(JB): 0.563
- Kurtosis: 1.738 Cond. No. 1.77e+06

--====Insights===--

--==== Interpretation of the regression output for avg units vs order year=====-

Summary of the model:

- **R-squared:** The R-squared value is 0.660. This means that approximately 66.0% of the variability in the average units sold can be explained by the linear regression model. In other words, the model is a reasonably good fit for the data.
- **Order_Year:** The coefficient for Order_Year is -65.5119. This implies that, on average, the number of units sold decreases by approximately 65.51 units each year. The negative sign indicates a decreasing trend.

P-values:

The p-value for the Order_Year coefficient is 0.014, which is less than 0.05. This suggests that the Order_Year variable is statistically significant in predicting the average units sold.

Overall Fit:

The model we used is good at explaining about 66% of the changes in the average units sold.

Yearly Change:

Every year, on average, there is a decrease of about 65 units in the number of items sold.

Conclusion:

- Over the years, the trend is showing a decline in the number of units sold.
- So, if this trend continues, we might sell fewer units in the future.
- The sellers might have to study customer patterns, market behaviour and trends to increase the average units sold in the upcoming years.

Overall Insights and recommendations:

- Croatia & UK are some of the countries with the lowest avg time to ship.
- Cosmetics has the highest avg profit whereas Fruits have the lowest. Steps can be taken to ensure that the profit for Fruits section improves.
- Every year on average, the model predicts that there would be a decrease of about 65 units in the number of items sold. If this trend persists there might be a decrease in the overall profit.
- The only department in which the avg profit is more(almost double) than avg cost was the Clothes department.
- Moreover, their avg days to ship is one of the lowest which makes them a sure shot “Winner” 🤩
- Sellers might have to prioritize on marketing and promotion strategies by introducing different sale seasons to boost profitability.
- The yearly average profit was the lowest in 2016 and the highest in 2012.
- The Avg profit between 2012 and 2013 has been the lowest which is -22.33%.
- Hungary has the highest avg days to ship followed by Georgia.
- The latter has to improve on the trade performance since its avg days to ship is one of the highest and avg profit is one of the lowest.
- Online sales have generated less profit than offline sales and steps can be taken to improve the latter’s performance.