Introduction:

The production of wine has widely impacted the global markets due to the demand and supply of different quality of wine in the world. Wine is consumed as drink in multiple regions in the world but especially European people manufacture grape wine. Wine is also imported to many countries due the taste and quality. Grapes are a multi-colored fruit that come in a few different colors like red, black, blue, dark blue, yellow, green, orange and shades. Interestingly enough, the grape varieties usually called "white" are not white as we commonly assume and have evolved from the purple ones. This is genetically linked to mutations in two regulatory genes turning off the synthesis of anthocyanin (the pigments of the deep-purple grapes) operation. Whereas the anthocyanin, together with other polyphenols that belong to the pigment family, are responsible for the color graduation from red to almost black wines. Normal clusters of grapes contain between 15 and 300 grapes, which are prelate spheroid shaped similar to an ellipsoid (Khan et.al, 2020).

The chemical composition of its raw form consist of at least 81% water, 18% carbohydrates, 1% protein and very little fat. A rough estimation of 69 calories are contained in a 100 gram size serving of fresh grape, which contains a moderate component of vitamin K (14% Daily Value).As Food and Agriculture Organization (FAO) states, about 75,866 sq. km of Earth's land is contributed by the grape cultivation only. The world grape production statistics indicate that 71% is for making wine, whereas 27% of grapes are eaten fresh in the market and 2% is dried to eat. Aside from this, a part of the production of grapes is utilized for grape juice that is often, reconstituted for canned fruits branded as "with no added sugar" and "100% natural." Of note, the area under vineyards is increasing annually, the average growth rate being given as 2% per year. The top five European wine producing countries are Italy, France, Spain, Germany and Portugal located) with 49.1 million hectolitres (mhl), 46.6 mhl, 40.7 mhl, 8.4 mhl and 6.4 mhl respectively (OIV, 2021). The aforementioned countries collectively represent the production of about the 58% of wine all over the world. Cumulatively, they export about 19 billion euros worth viticulture products, emphasizing the economic importance of wine for these areas. In consideration of the fact that wine making is one of the most critical parts of the economies of the chosen countries, further research was undertaken to assess their performance. Besides, France, Germany and Spain also have a link called Moselle valley which is a common wine region providing more evidences to enlist them in the research. The analysis of this study focuses on the main wine regions which are covered in the period under study that is from 1986 to 2015.

Many mechanized harvesting methods suffer to detect and sort the sunburned and shriveled berries (SSB), which end up in the harvest batch overwhelmed by underripe and damaged grapes. In Tuscany, Italy, until the 1990s, the conventional harvesting method was manual picking that was followed with mechanical destemming which didn't include post-harvest sorting. Consequently, the introduction of a re-planting programme in the 1990s that led to the creation of many new farms made mechanical harvesting a more popular practice. In spite of this style, mechanical harvesting is infrequently found, because it is expensive to use harvesters and the average area in estates is about nine hectares on average. For example, there is the question of how to handle the large amounts of both SSB and MOG when they are present among the grapes that are going to be harvested (Yang et.al, 2022). While today's harvesters are fitted with onboard MOG removal units since the 1970. Some materials are still likely to make their way into the winemaking process. Their shape, size or frictional interactions can potentially influence the chemical and sensory properties of the resulting wine (Parenti et.al, 2015) .

Tuscan vintners lately have joined those who are working with post-harvesting sorting. The most popular technique utilizes number of vibrating sorting tables with conveyor belts to enable visual inspection which sorts substandard berries and extraneous masses manually. Finally, automatic grape sorting machines starts to become more popular. For example, densimetric sorting machines not only separate berries with a higher density, but also help remove berries with a lower density after stem removal. Although these machines have proved their usefulness they are pricier than vibrating sorting tables, and they are not so often adopted in Italy as of now.

Task 1:

The dataset has been designed by taking all the preprocessing steps with the usage of excel ETL techniques. Initially, an existing dataset is taken from a repository of github then columns are distributed to different regions as France, Germany, and Spain. Various factors of grapes have been included. Quantitative data encompasses numerical information, categorized into two main types that are discrete and continuous, corresponding to ranges of values which are countable and can span any value which falls within a certain range respectively. However, qualitative data is used to evaluate attributes or components; it faces quantification issues, for example, in customer reviews. Structured data is particularly organized in predefined tables or databases making analysis and accessibility without hassle. However, disorganized data does not have a predefined organization and as a result, has to be processed which includes the case of text documents. Semi-structured data encompasses this where data is semi-organized without rigid structured.

Data Preprocessing

Data preprocessing means carrying out a number of essential operations which tend to improve the quality of the dataset before it gets on processing. Usually this process starts with data cleaning where abnormalities such as duplicates, outliers and blank values are detected, rectified and filled by techniques like imputation and transformation. Continuously, the need for data integration comes that allows to combine the data from multiple sources in one cohesive one, and therefore accurate attention is needed to deal with problems of formatting, structures, and semantic. After the integration of data, comes the transformation stage, in which the data undergo the processes of standardization, normalization, and discretization, with the objective of uniformity and compatibility for analysis. Together with data reduction strategies that could be used to stream the datasets by eliminating the less important information at the loss of key parameters through methods like feature selection as well as extraction. These steps that are related to preprocessing are overall the basis for perfect data mining, leading to integral insights that are then utilized to build sound decision making on data of high quality.

Applied Preprocessing:

1. Cells with value of 0 the cells remain unaltered if they follow consistent patterns in data rows.
2. It is crucial to address any null or missing values in the dataset for the analysis to be done correctly. One of the strategy the missing values can be handled is by replacing them with the average values that had been calculated from the adjacent data points. The repetitive values and the outliers have been identified and adjusted to align to the value range of the whole dataset.
3. The dataset inconsistencies are preserved too as they are going to be illustrated in the task 3 dashboard design. In the process of determining duplicate values as purposeful method is used to define their contribution. For instance, repetition of duplicates within ten to fifteen copies and it is allowed to keep them because it is a common thing that it can be found in the text. On the instances of overabundance such duplicates occurring excessively from ten to twenty times and differing sharply from the average values, need to be corrected and brought to adjustment with the average dataset.

Country column addition:

Treatmen	Replicatio	Cluster we	Berry wei	TSS	TA	Firmness	Multiplex	Multiplex (	Convex Area	Country
1.000	1.000	484.210	8.095	22.400	0.373	4.745	3.750	0.305	0.100	France
1.000	2.000	592.390	9.715	21.200	0.354	5.805	4.735	0.486	0.100	France
1.000	3.000	605.740	9.055	19.800	0.436	5.936	3.610	0.258	0.000	France
1.000	4.000	619.480	12.850	19.800	0.332	6.071	3.798	0.217	0.100	France
1.000	5.000	629.730	9.635	18.900	0.368	6.171	3.336	0.202	0.000	France
1.000	6.000	618.180	9.480	19.800	0.373	6.058	3.856	0.350	0.000	France
1.000	7.000	571.150	8.025	20.400	0.399	5.597	4.727	0.549	0.100	France
1.000	8.000	645.430	8.095	21.200	0.405	6.325	4.002	0.505	0.100	France
1.000	9.000	580.270	9.405	19.700	0.349	5.687	4.379	0.276	0.100	France
1.000	10.000	647.860	7.770	20.600	0.435	6.349	2.804	0.214	0.100	France
2.000	1.000	664.930	8.350	23.100	0.396	6.516	4.023	0.412	0.100	France
2.000	2.000	671.740	10.265	19.800	0.426	6.583	3.935	0.279	0.100	France
2.000	3.000	650.830	9.720	20.000	0.426	6.378	3.814	0.284	0.000	France
2.000	4.000	571.540	9.180	21.300	0.397	5.601	4.532	0.286	0.100	France
2.000	5.000	660.010	8.715	20.800	0.387	6.468	4.846	0.412	0.100	France
2.000	6.000	657.260	11.180	19.100	0.394	6.441	4.603	0.326	0.000	France
2.000	7.000	565.110	8.685	20.000	0.406	5.538	3.151	0.204	0.100	France
2.000	8.000	637.180	10.185	19.600	0.324	6.244	4.041	0.340	0.100	France
2.000	9.000	664.060	11.480	21.300	0.447	6.508	4.137	0.505	0.000	France
2.000	10.000	619.250	9.095	22.100	0.312	6.069	2.561	0.264	0.100	France
3.000	1.000	639.710	9.655	20.100	0.460	6.269	4.672	0.461	0.000	France
3.000	2.000	519.580	8.880	19.300	0.406	5.092	2.657	0.250	0.000	France

Duplication Columns

71	8.000	10.000	672.690	12.280	20.400	0.451
72	9.000	1.000	657.550	6.855	22.300	0.531
73	9.000	2.000	469.300	10.280	20.700	0.498
74	9.000	3.000	591.420	10.745	20.600	0.419
75	9.000	4.000	689.280	8.445	19.700	0.465
76	9.000	5.000	570.190	10.315	19.700	0.393
77	9.000	6.000	642.660	9.200	19.700	0.466
78	9.000	7.000	667.740	6.325	20.700	0.528
79	9.000	8.000	672.660	10.805	19.700	0.516
80	9.000	9.000	567.060	8.955	20.000	0.438
81	9.000	10.000	557.300	9.885	21.200	0.441
82	10.000	1.000	642.750	9.240	21.700	0.444
83	10.000	2.000	591.030	10.025	20.300	0.408
84	10.000	3.000	606.350	10.295	20.200	0.450
85	10.000	4.000	471.900	10.495	20.700	0.397
86	10.000	5.000	581.010	9.665	20.400	0.436
87	10.000	6.000	583.650	7.455	21.700	0.355
88	10.000	7.000	596.890	9.805	21.100	0.453

Germany Column:

21.100	0.505	3.244	4.117	0.540	1.040	Germany
20.700	0.582	2.411	4.108	0.385	0.913	Germany
19.400	0.594	5.504	3.709	0.163	1.174	Germany
18.700	0.450	4.685	3.916	0.425	1.062	Germany
19.800	0.575	3.312	3.889	0.313	1.129	Germany
21.000	0.549	1.685	4.518	0.284	1.078	Germany
21.800	0.537	2.002	3.467	0.384	1.060	Germany
19.800	0.535	3.951	4.707	0.449	1.384	Germany
20.400	0.497	4.503	3.218	0.197	1.103	Germany
20.700	0.544	3.786	4.274	0.442	1.104	Germany
20.200	0.573	3.774	2.649	0.267	1.303	Germany
20.000	0.583	3.272	4.456	0.410	1.303	Germany
20.000	0.513	2.798	4.586	0.439	1.202	Germany
20.700	0.530	3.207	3.641	0.435	1.281	Germany
19.300	0.492	3.890	5.273	0.403	1.741	Germany
18.900	0.507	3.844	4.838	0.461	1.389	Germany

Spain data addition:

18.500	0.469	4.170	4.540	0.418	1.370	Spain
20.700	0.435	2.230	4.247	0.362	1.502	Spain
20.400	0.400	3.467	3.676	0.305	1.891	Spain
19.200	0.313	2.462	3.756	0.462	1.322	Spain
19.700	0.333	2.915	4.356	0.622	2.253	Spain
21.500	0.345	1.910	3.809	0.348	1.284	Spain
21.100	0.384	3.287	3.712	0.418	1.358	Spain
19.300	0.393	3.178	4.680	0.579	1.509	Spain
19.600	0.333	3.990	3.232	0.131	1.269	Spain
22.400	0.394	2.282	3.767	0.453	0.868	Spain
21.200	0.439	1.873	3.351	0.197	1.605	Spain
20.200	0.438	2.071	4.244	0.349	1.156	Spain
20.000	0.474	3.536	3.727	0.407	0.755	Spain
20.500	0.404	2.817	4.756	0.592	0.867	Spain
19.200	0.390	2.253	5.436	0.511	1.446	Spain
20.400	0.344	1.657	4.675	0.485	1.495	Spain
20.400	0.508	2.090	3.328	0.335	0.963	Spain
21.100	0.508	1.405	3.898	0.551	0.968	Spain
19.200	0.327	1.969	5.380	0.822	1.282	Spain
19.700	0.363	2.805	3.689	0.273	1.603	Spain
20.200	0.426	3.827	3.845	0.346	0.936	Spain
22.300	0.420	2.029	5.142	0.584	1.448	Spain
19.300	0.358	3.995	3.571	0.318	1.450	Spain

In the above steps the journey of data preprocessing such as a series of essential steps has been systematically executed culminating in a refined dataset ready for insightful analysis. The initial phase involved comprehensive data collection from diverse sources such as surveys sensors and databases. Following the data collection an intensive data cleaning phase was undertaken to rectify errors and inconsistencies and address missing values. This step was crucial in enhancing the overall quality and reliability of the dataset. The subsequent stages of data transformation were instrumental in shaping the data for efficient analysis. Through careful reshaping by including unit conversion and variable creation the data became more amenable to exploration. A critical data reduction phase is implemented to streamline the dataset by focusing on pertinent information and reducing dimensionality. This involved summarizing text data and retaining only relevant features by contributing to a more concise and manageable dataset. With each preprocessing step the data underwent a refining process addressing intricacies and optimizing its structure for subsequent analytical endeavors. The comprehensive nature of these preprocessing steps has successfully laid the groundwork presenting a clean integrated and transformed dataset ready for meaningful analysis and interpretation.

Task 2:

Descriptive statistics serves as the foundational cornerstone of statistical analysis, offering a comprehensive and concise summary of a dataset's main characteristics. It provides a clear picture of the data's overall composition, enabling the identification of potential patterns, anomalies, and trends. Commonly employed in statistical analysis are various steps of descriptive statistics.

1. Mean: This is the indicator of the average data value. It is calculated by adding all the data values and dividing the sum by the total number of data values. The central mean give us the possibility to know as representative the points of data.

2. Median: Considering that the middle value rank is the same whether data points are lined up in the ascending order or descending order, the median appears very stable and is a good measure of central tendency in fact.

3. Mode: This means that the value that repeats the most in the data is the one that takes place most often. Consequently, the data contains the most common observation.

4. Range: The range is spread between the most and the least value in the dataset. Ranging from the extremes gives a simple assessment.

5. Variance: To be determined as the average of the square of the deviations of the data points from the mean, variance provides the knowledge about how spread out the data is around the mean.

6. Standard Deviation: This is the square root of the variance and it is measured in the units same as the data. It is easier to interpret as an alternative to variance because it is a more direct measure of spread.

7. Skewness: The skewness measures are the symmetry of the data distribution. A positive skew stands for the right tail, and a negative skew for the left tail, as the divergence from a normal distribution is higher for the right and lower for the left tail.

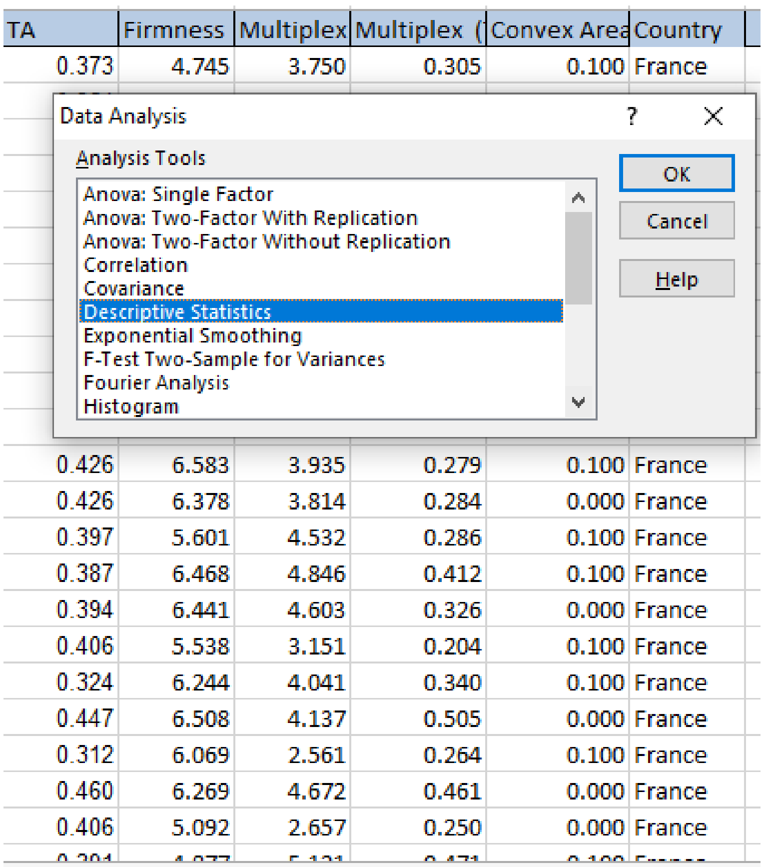
8. Kurtosis: We denote by the peak of the data distribution the skewness of a normal “bell-shaped” curve which shows the sharpness or flatness of the peaks and tails.

9. Percentiles: Through splitting the data into hundred even divisions, percentiles indicate the amount of data which is below a specific percentage.

10. Frequency Distributions: These are frequency of data points within certain ranges of values commonly represented by histograms.

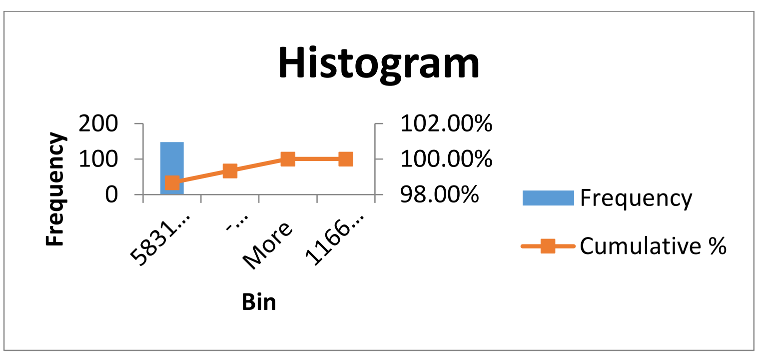
11. Boxplots: Visualizing the distribution of data, a box plot illustrates the median, quartiles, and any outliers.

Applying Descriptive Statistics:

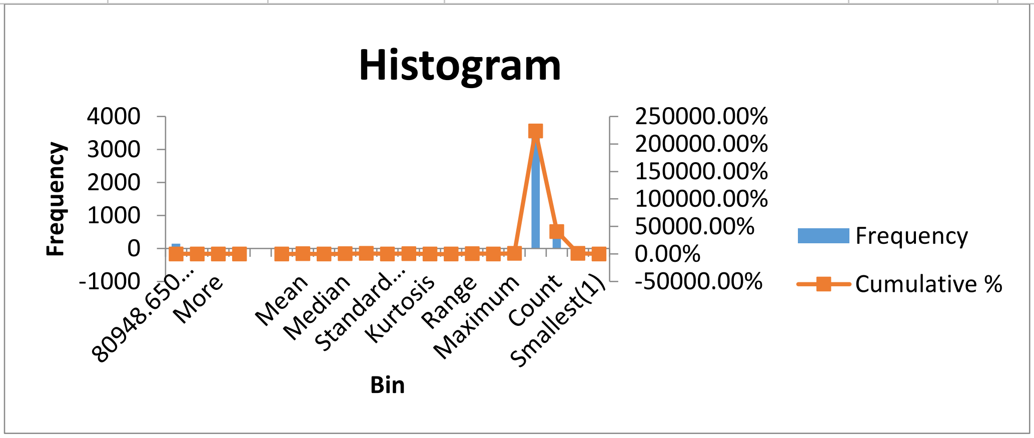


TA	Firmness	Multiplex	Multiplex (	Convex Area	Country
0.373	4.745	3.750	0.305	0.100	France
0.426	6.583	3.935	0.279	0.100	France
0.426	6.378	3.814	0.284	0.000	France
0.397	5.601	4.532	0.286	0.100	France
0.387	6.468	4.846	0.412	0.100	France
0.394	6.441	4.603	0.326	0.000	France
0.406	5.538	3.151	0.204	0.100	France
0.324	6.244	4.041	0.340	0.100	France
0.447	6.508	4.137	0.505	0.000	France
0.312	6.069	2.561	0.264	0.100	France
0.460	6.269	4.672	0.461	0.000	France
0.406	5.092	2.657	0.250	0.000	France

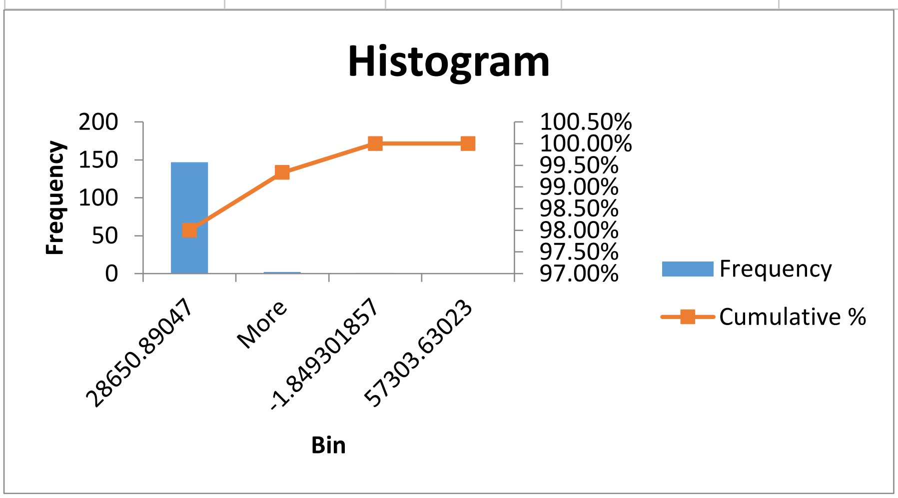
DescriptiveStatisticsFrance							
	Treatment	Replication	Cluster weight (T0)	Berry weight	TSS	TA	Firmness (T0)
Mean	8.172414	5.5	603.2781	9.517724	20.3114	0.41664	5.886214
Standard Error	0.258061	0.1689577	3.799412	0.093788	0.056438	0.00281	0.035975
Median	8	5.5	618.19	9.45	20.22206	0.419	5.954697
Mode	1	1	607.71	9.45	20	0.396	5.955558
Standard Deviation	4.394625	2.8772464	64.70166	1.597148	0.961111	0.04786	0.612624
Sample Variance	19.31273	8.2785467	4186.305	2.550882	0.923734	0.002291	0.375309
Kurtosis	-1.12928	-1.224637	0.247257	0.293645	0.553489	-0.03223	0.372409
Skewness	-0.02292	0	-0.791201	0.291744	0.169546	0.322321	-0.85529
Range	15	9	337.49	9.095	6.4	0.258	2.957248
Minimum	1	1	415.25	5.735	17.5	0.288	4.06945
Maximum	16	10	752.74	14.83	23.9	0.546	7.026698
Sum	2370	1595	174950.6	2760.14	5890.305	120.8255	1707.002
Count	290	290	290	290	290	290	290
Largest(1)	16	10	752.74	14.83	23.9	0.546	7.026698
Smallest(1)	1	1	415.25	5.735	17.5	0.288	4.06945

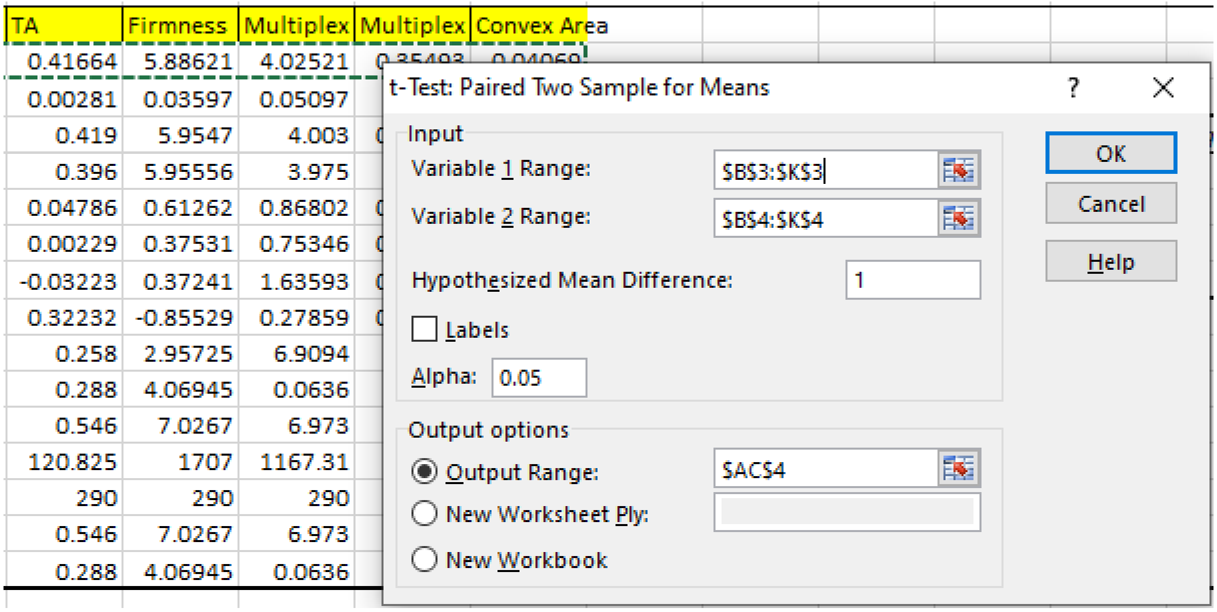


DescriptiveStatisticGermany							
	Treatment	Replication	Cluster weight (T0)	Berry weight	TSS	TA	Firmness (T0)
Mean	8.9511	5.4718826	593.7614	9.373533	19.69835	0.46284	2.761877
Standard Error	0.214601	0.1430327	3.116939	0.080309	0.080563	0.004259	0.044195
Median	9	5	607.24	9.4	20.1	0.468	2.791241
Mode	7	10	586.29	9.45	20.7	0.492	3.243908
Standard Deviation	4.340027	2.8926575	63.03619	1.62414	1.629289	0.086128	0.893796
Sample Variance	18.83584	8.3674673	3973.561	2.637829	2.654582	0.007418	0.798871
Kurtosis	-1.02892	-1.234053	0.083531	0.323582	0.792406	-0.89864	-0.25779
Skewness	-0.17905	0.0186329	-0.711519	0.18903	-0.82668	0.095634	0.284729
Range	15	9	332.12	9.62	11.2	0.395	5.000519
Minimum	1	1	409.67	4.595	13.1	0.289	0.503593
Maximum	16	10	741.79	14.215	24.3	0.684	5.504111
Sum	3661	2238	242848.4	3833.775	8056.625	189.3016	1129.608
Count	409	409	409	409	409	409	409
Largest(1)	16	10	741.79	14.215	24.3	0.684	5.504111
Smallest(1)	1	1	409.67	4.595	13.1	0.289	0.503593



Descriptive Statistic Spain							
	Treatment	Replication	Cluster weight (T0)	Berry weight	TSS	TA	Firmness (T0)
Mean	9.991803	5.4836066	352.2802	7.889057	17.20123	2.542029	7.009265
Standard Error	0.282025	0.1831522	18.59476	0.192722	0.279398	0.198628	0.342403
Median	11	5	524.24	8.7275	19.25	0.554	3.648491
Mode	13	3	13	6	13	6	13
Standard Deviation	4.405376	2.8609291	290.4594	3.010413	4.364338	3.102669	5.348499
Sample Variance	19.40734	8.1849153	84366.64	9.062586	19.04745	9.626556	28.60644
Kurtosis	-0.89002	-1.208104	-1.849302	-0.19596	2.474609	-0.0348	-1.56654
Skewness	-0.58878	0.0139659	-0.264997	-0.69004	-1.34387	1.199299	0.499911
Range	15	9	708.02	13.215	22.3	9.711	15
Minimum	1	1	1	1	1	0.289	1
Maximum	16	10	709.02	14.215	23.3	10	16
Sum	2438	1338	85956.37	1924.93	4197.1	620.255	1710.261
Count	244	244	244	244	244	244	244
Largest(1)	16	10	709.02	14.215	23.3	10	16
Smallest(1)	1	1	1	1	1	0.289	1





France

t-Test: Paired Two Sample for Means

	Major	*Minor*
Mean	65.75032831	0.447853061
Variance	35707.18901	1.393435957
Observations	10	10
Pearson Correlation	0.997807063	
Hypothesized Mean Difference	1	
df	9	
t Stat	1.082842632	
P(T<=t) one-tail	0.153517479	
t Critical one-tail	1.833112933	
P(T<=t) two-tail	0.307034957	
t Critical two-tail	2.262157163	

Germany

t-Test: Paired Two Sample for Means		
	Major	*Minor*
Mean	64.66786042	0.376427193
Variance	34594.09962	0.931380502
Observations	10	10
Pearson Correlation	0.998244311	
Hypothesized Mean Difference	1	
df	9	
t Stat	1.081681066	
P(T<=t) one-tail	0.153761961	
t Critical one-tail	1.833112933	
P(T<=t) two-tail	0.307523921	
t Critical two-tail	2.262157163	

Spain

t-Test: Paired Two Sample for Means		
	Major	*Minor*
Mean	41.58227143	2.079172461
Variance	11935.21692	33.68212677
Observations	10	10
Pearson Correlation	0.999329556	
Hypothesized Mean Difference	1	
df	9	
t Stat	1.176981999	

P(T<=t) one-tail	0.134695628
t Critical one-tail	1.833112933
P(T<=t) two-tail	0.269391255
t Critical two-tail	2.262157163

The statistical analysis of data shows the variety of grapes based on histogram visualization, descriptive statistics, and inferential hypothesis tests suggest the varieties of grapes in France have high quality attributes as per the dataset analyzed. While the variety of France regions is different. It is suggested that Wines that are produced from the grapes of France can be sold high prices while Spanish brand should be lower in terms of price than the brand Germany. The Price of the wine from France region grapes should be higher than both German and Spanish. The research undertakes giving the probable reasons that explain the quality expositions and value features of wine in Germany and Spain. In the German case, the rivalry of wineries between the private and the cooperative wineries emphasizes the point that quality of grapes and their high market reputation dictates the fate of the wineries. A major issue that cooperative wineries may confront is maintaining a consistent level of quality, as this can be made more difficult by their reliance on grapes brought in from individual growers, contrasted with the greater production control over their supply chain. In turn, this may result in higher quality wines from private wineries. The arise respect for wines quality brings dramatic dynamics and policy issues respectively.

On the other hand, the fact that climatic factors like temperature and the amount of rainfall have a direct impact on the maturity of grapes and the quality of wines demonstrate how complex is the connection between environment and viticulture as farming activity. It becomes apparent that temperature being somehow connected with wine production and, therefore, specific environmental conditions may elevate the quality of grapes, while the rainfall that falls during the bloom time is also able to perform a positive role. Understanding these relationships is an integral component to cultivating grapes for production and wine quality in specific regions. The Spanish study casts a spot light on the intricate relationship between various winemaking practices, grape quality, and wine taste. The main focus of the research is on the evaluation of different protection methods to understand how this affects the yields and nutrient contents. Consequently, the researchers try to point out ways of improving viticulture practices and the quality of the grapes. The article pinpoints on the necessity for additional study in order to fully illuminate what effect has the noticeable higher magnesium content on vine physiology and wine quality. Researchers can move a step further on nutrition and nutrients combinations that enhance disease tolerance by investigating how they work in the vineyards. This way, they can come up with more effective and sustainable disease management strategies. Current investigations support developing the explanation of the influence of some factors on the quality and reputation of Germany and Spain wines. Through clarifying the mutually influencing variables that affect the environment, agriculture, and management factors, researchers could inform practices and produce wine which will not just be of best quality but will be sustainable as well.

Excel Formula:

STANDARD DEVIATION function = STDEV.P(number1,[number2],…)

VARIANCE function = VAR(number1, [number2], …)

QUARTILES function = QUARTILE(array, quart)

CORRELATION function = CORRE(array1, array2)

COUNT function = COUNT(value1, [value2], …)

COUNTA function = COUNTA(value1, [value2], …)

AVERAGE function = AVERAGE(number1, [number2], …)

AVERAGEIF function = AVERAGEIF(range, criteria, [average_range])

MEDIAN function = MEDIAN(number1, [number2], …)

MODE function = MODE.SNGL(number1,[number2],…)

Task 3:

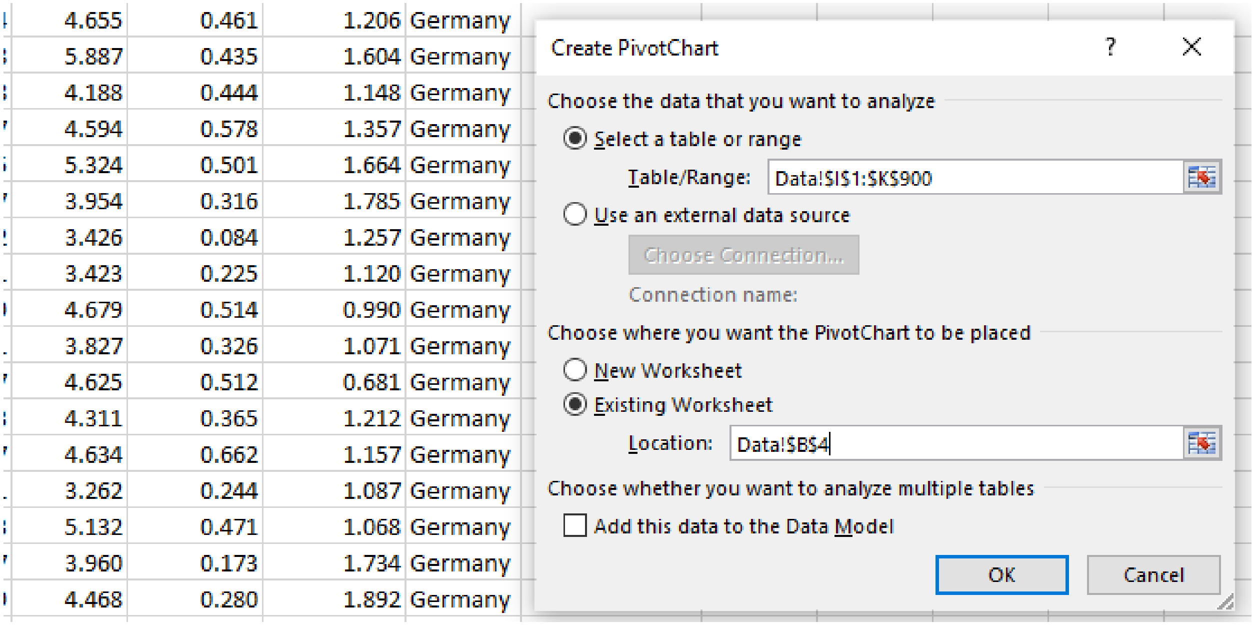
The analysis of data from both task 1 and 2 is visualized into this part to show the results that will help in decision making process to the stockholders. This analysis aimed to help all the parameters of grapes that support in the quality and prices selection of wine.

Excel Data Analysis

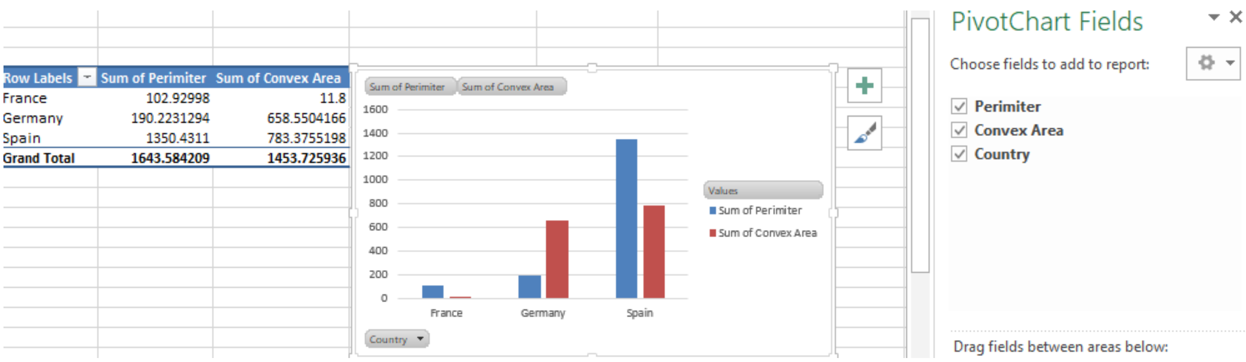
The first step in excel data analysis is to import the data into the spreadsheet. These data can be sourced from different sources such as database, CSV files, or even excel sheets of other people. Excel applications work with data handling such as sorting, filtering and summarizing. The other significant area in the data analysis in Excel is the utilization of formulae and functions that allow you to do the calculations and obtain new insights from the data. Excel's set of functions is rich, including the most basic arithmetic functions, statistics, and a vast number of functions useful for complex analysis operations.

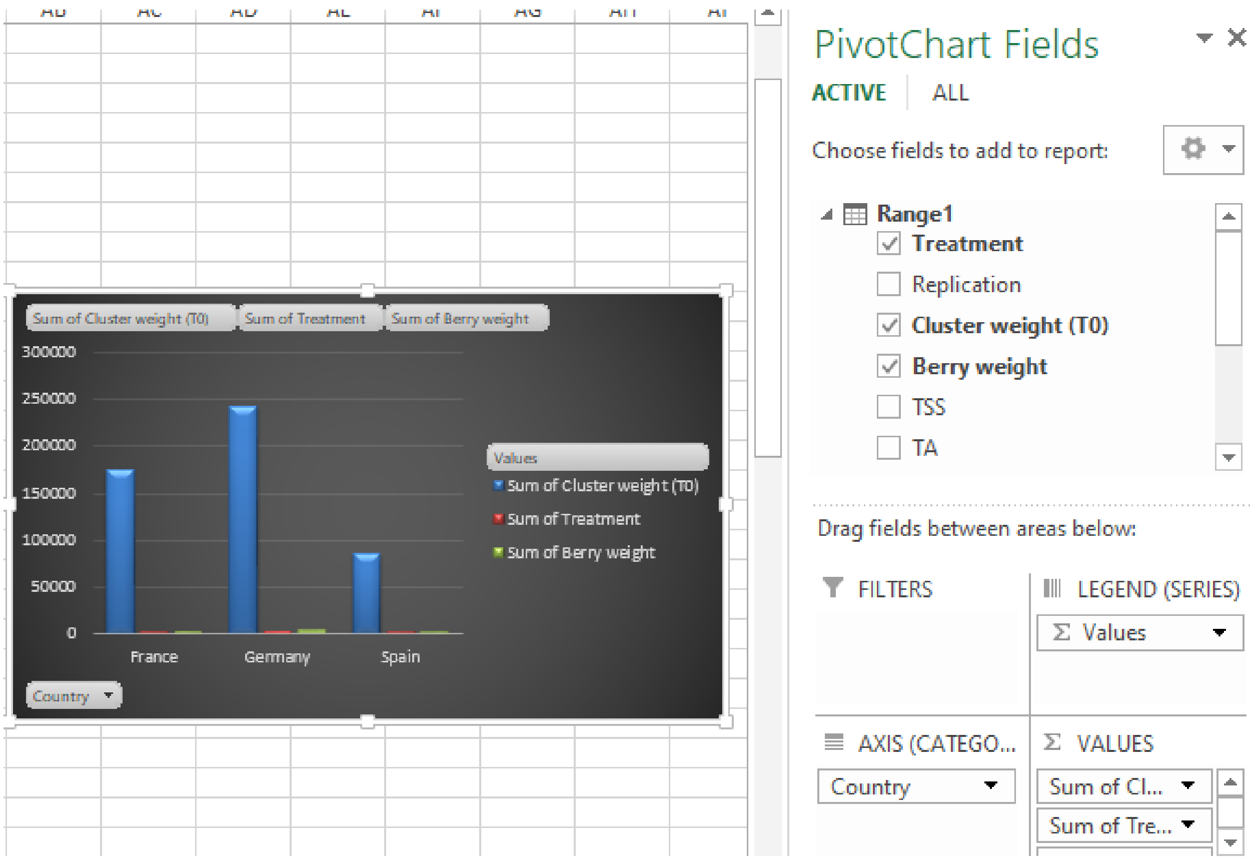
The computing functions can be used to estimate averages, standard deviations, correlations, and other statistical indicators that help professionals to interpret underlying patterns in the data. The capabilities of Excel enable data analysts to obtain these variables from datasets of different sizes and levels of complexity to provide insightful analysis. Excel being the tool with the largest number of functions, calculations and features in its framework, analysts can perform really wide scope of analysis from basic to complex statistics. Spreadsheets in Excel organize data in a structured manner and makes it easier to carry out manipulations and data transformations into informative stories. Some of the tools that you can use to represent the data in graphics such as pivot tables, charts, and graphs will help you to identify trends, distributions, and relationships among the variables, thus, your stakeholders will absorb it at a glance. First, the Excel's validation and filtering abilities increase data's quality and help to concentrate well on the subsets of data that are actually going to be analyzed in deeper details. Secondly, Excel's ability to interactively link other Microsoft Office applications and read/write from other data sources ensure the easy import/export of data and teamwork among colleagues. Generally, Excel has a wide range of useful functions for data analytics that enable its users to obtain actionable intelligence based on which right decisions can be unambiguously made in many areas and referring to few industries.

Create Pivot chart:

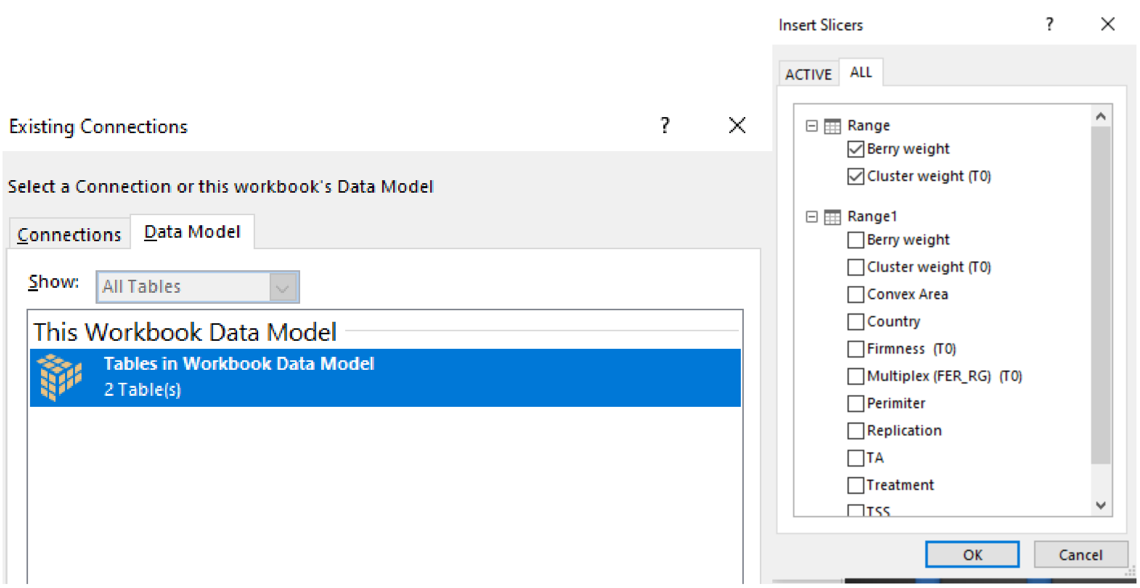


Generated Pivot chart



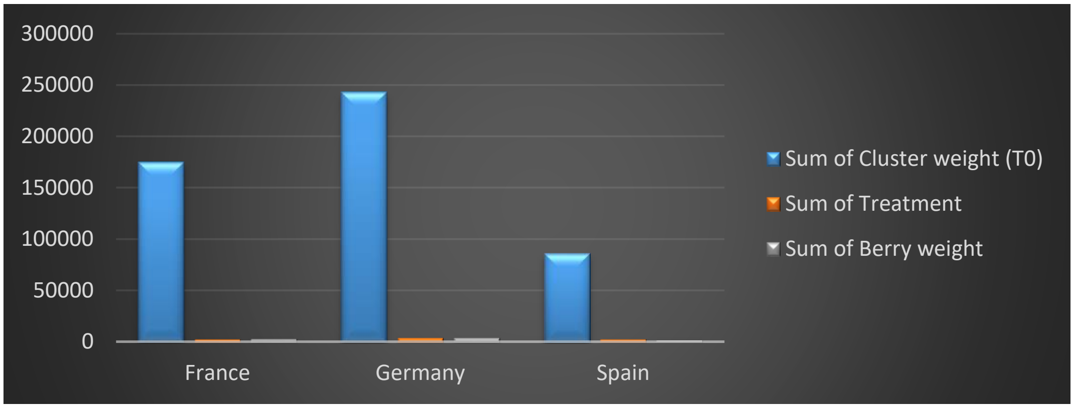

Sum of Cluster weight (T0)
Sum of Treatment
Sum of Berry weight
300000
250000
200000
150000
100000
50000
0
France
Germany
Spain
Values
Sum of Cluster weight (T0)
Sum of Treatment
Sum of Berry weight
Country
PivotChart Fields
ACTIVE
ALL
Choose fields to add to report:
Range1
Treatment
Replication
Cluster weight (T0)
Berry weight
TSS
TA
Drag fields between areas below:
FILTERS
LEGEND (SERIES)
Values
AXIS (CATEGO...
VALUES
Country
Sum of Cl...
Sum of Tre...

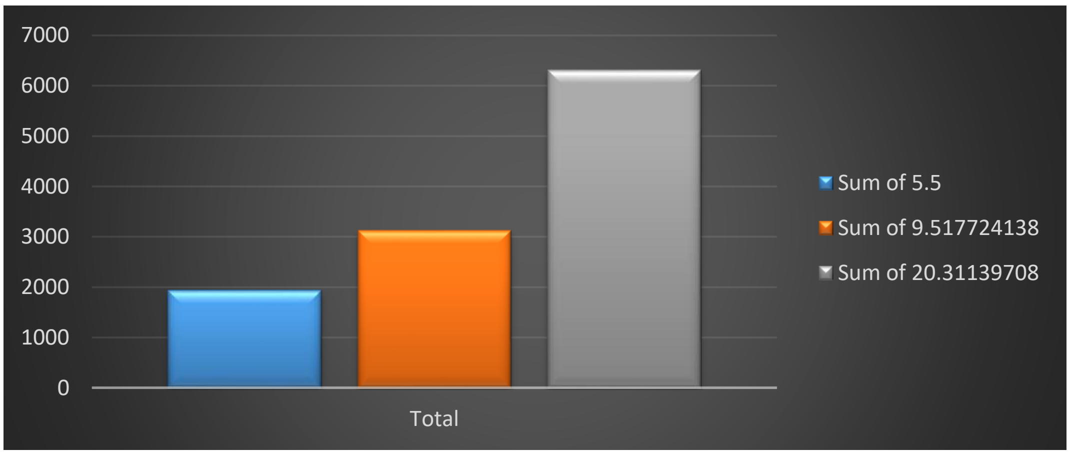
Slicers:

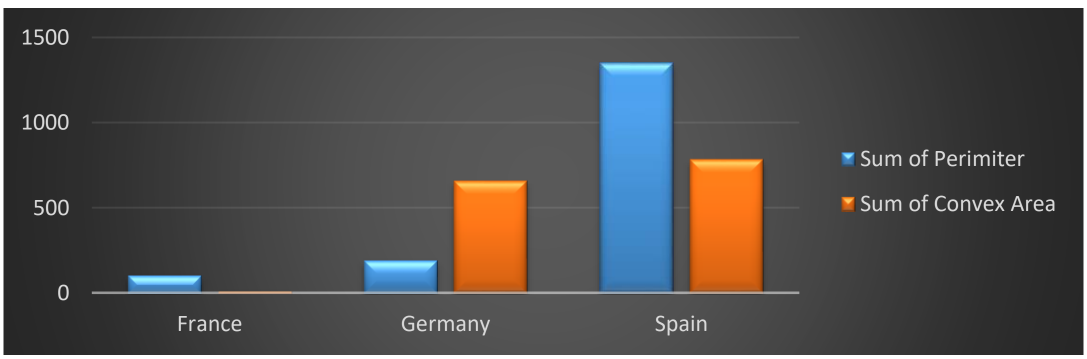

Existing Connections
Select a Connection or this workbook's Data Model
Connections
Data Model
Show:
All Tables
This Workbook Data Model
Tables in Workbook Data Model
2 Table(s)
Insert Slicers
ACTIVE
ALL
Range
Berry weight
Cluster weight (T0)
Range1
Berry weight
Cluster weight (T0)
Convex Area
Country
Firmness (T0)
Multiplex (FER_RG) (T0)
Perimiter
Replication
TA
Treatment
TSS
OK
Cancel

Dashboard Charts

Following charts are the generated pivot charts from the analysis. It shows the data of grape harvest of the France, Germany, and Spain regions.







Full Dashboard Design

Row Labels	Sum of Perimiter	Sum of Convex Area
France	102.92998	11.8
Germany	190.2231294	658.5504166
Spain	1350.4311	783.3755198
Grand Total	1643.584209	1453.725936

Mean	(All)	
Sum of 5.5	Sum of 9.517724138	Sum of 20.31139708
1932.600114	3124.092207	6312.391527

In Europe, wine quality and consumption are inseparable from the majestic cultural heritage as well as the countless winemaking traditions rooted in this immense continent. European nations, such as France, Italy, and Spain, are well known for having some of the world's best wine collections, and each variety has its specific features and flavor. The terroir concept, meaning the particular environmental elements that affect vine growing, is another key to understanding European wines taste and difference. This preference for terroir, developed over the centuries, as well as the predominance of high-quality over quantity is what enables Europe to maintain its status as the wine industry leader in the world market.

In the area of food consumption the wine keeps a distinguished place in the European culture which is usually taken together with meals, friend gatherings, and holidays. Wine is not only a beverage for many European countries, but it is also an issue of tradition, history and identity. Even if wine consumption in Europe experiences a few ebbs and flows lately, in some of the traditional wine-drinking countries, overall, wine is an important part of the European life style

One of the most significant influences of the European Union on wine industry is regulation of it within the member states by means of different initiatives aimed at guaranteeing wine quality, protecting geographical indications, and supporting winemaking that is eco-friendly. The European wine industry keeps changing with time, with the idea of sustainability, innovation and readiness to new consumer expectations becoming its major pillars.

Applying the grape harvest data to identify the wines prices of three different countries: France, Germany and Spain, France demonstrates her pre-eminence i in both quality and price. The data visualisation tells a really efficient story that stresses wooden way higher corresponding quality of the French wines than of the German and Spanish wines, which is brought about by huge sums these wines cost on the market compared to those of Germany and Spain. Through careful watch of the grape harvest data and the corresponding wine prices, the stakeholders gain significant knowledge of market trends and consumers' wine choices within the wine industry. Moreover, the visualization demonstrates the prevalence of French wines. On the other hand, the visualization presents the story of not only the position of the French wines as being superior but also the story of looking into the possible factors that contribute to the high pricing of the wines, such as renowned vineyard regions, traditional winemaking techniques, and established brand reputation. The thorough knowledge of the wine market landscape provides the stakeholders with the necessary ingredients to make educated responsibilities on product positioning, investment strategy and product differentiation so as to take advantage of the huge opportunities presented by the sector. As a result, the implementation of data visualization in the matter of grape harvest is providing invaluable information with a clear view of business strategy, to enable the company in achieving market share in the international market.

Conclusion :

Assessing the correlation between grape harvest data and wine value in French, German and Spanish markets, has generated a pool of information on what affects market's perception of grape quality and by extension wine price. The grape harvest data analysis shows a remarkable situation of France as a pioneer: its grapes are superior to those of Germany and Spain in quality.

A very important outcome is that the freshness of grapes picked in France dictates wine cost in that area. Premium demand for and higher price of French wine are two implications of the better quality of French grapes than those grown in other places. A number of elements are responsible for this situation, like the centuries-long and well-known French wine-making tradition, suitable weather conditions grapes flourishing, as well as the established quality control system in the French winemaking industry. To the contrary, Germany and Spain produce, of course, noteworthy wines as well, but their vineyards do not reach the same level of high quality consistency in the production and quality of grapes as France. Therefore, a bottle of wine from those regions often goes for less than a French equivalent. Disregarding their low average cost, wines from Germany and Spain are still worth the money and have great general taste as well. They are also unique and diverse due to regional particularities. In conclusion, the findings from the data of grape harvest and wine samples indicate the markedly role of grapes in prices and the estimated worth of French, German and Spanish wine. If France becomes the world's leader in grape and wine quality, and also, Germany and Spain can play in their own way with their unique and outstanding offers. Through adopting data visualization skills stakeholders will be able to derive crucial trends and preferences information of their customers and utilize the data to make the wine industry more effective.

References

Khan, N., Fahad, S., Naushad, M. and Faisal, S., 2020. Grape production critical review in the world. *Available at SSRN 3595842*.

Yang, C., Menz, C., Fraga, H., Costafreda-Aumedes, S., Leolini, L., Ramos, M.C., Molitor, D., van Leeuwen, C. and Santos, J.A., 2022. Assessing the grapevine crop water stress indicator over the flowering-veraison phase and the potential yield lose rate in important European wine regions. *Agricultural Water Management*, *261*, p.107349.

Parenti, A., Spugnoli, P., Masella, P., Guerrini, L., Benedettelli, S. and Di Blasi, S., 2015. Comparison of grape harvesting and sorting methods on factors affecting the must quality. *Journal of Agricultural Engineering*, *46*, pp.19-22.