Bucharest Academy of Economics Studies

Faculty of Cybernetics, Statistics and Economic Informatics



**Software Packages Project**

Festival Tickets Sales

Students:

Barbu Laura Andreea

Grecu Andy

Bucharest

2021

**Chapter 1. Excel**

**1. Financial problems**

The problem: The company requires a deposit in order to finance the next festival in Caraib Island that will commence after two years. The initial deposit is 2.500.000euros, while the monthly deposits will amount to 65.000euro/Month. The bank pays 2.15% interest. What will be the final balance?

The solution: We are going to use the **FV** function in order to find the value of the balance after 2 years.

=FV(B11/12, B13\*12, -B12, -B10, 1)

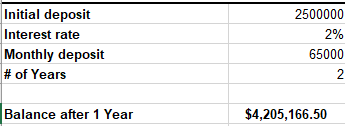


Fig 1.1 FV output

The problem: The company should organize a new festival during the winter season that will generate a profit of 1.375.000 euro. For this, the firm will ask for  a credit, the bank offering an interest rate of 10.35% per annum. Which is the maximum acquisition price if the festival liabilities among organization debts have to be paid out of the profits it brings, operating costs not being considered?

The solution: We are going to use the **PV** function in order to find the maximum acquisition price.

=PV(B3/12, B4\*12, -B2/12)

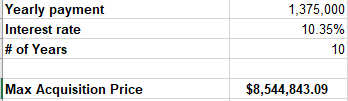


Fig 1.2 PV output

The problem: The company invests 4.520.000USD into a fund and receives the following yearly incomes: 755.000 euro,  1.245.250 euro, 355.120 euro. Supposing a 6% inflation rate, what will be the net value of the investment?

The solution:We are going to use the **NPV** function in order to find the net value of the current investment.

=NPV(B23,B19,B20,B21,B22)

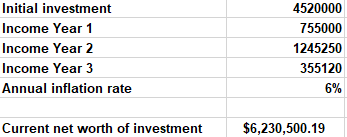
****

Fig 1.3 NPV output

**2. Lookup function**

The problem: In order to keep in evidence each day of ticket sales, the company placed in a new worksheet all the ticket sales on each day for each festival. Because new entries can appear in the future, the company wants to centralize all ticket sales from a specific date in a new list.

The solution: We use the compound function:

=IFERROR(INDEX($B$2:$B$1401,SMALL(IF($H$6=$A$2:$A$1401, ROW($A$2:$A$1401)-ROW($A$2)+1), ROW(1:1))),"" ) to obtain the following results:

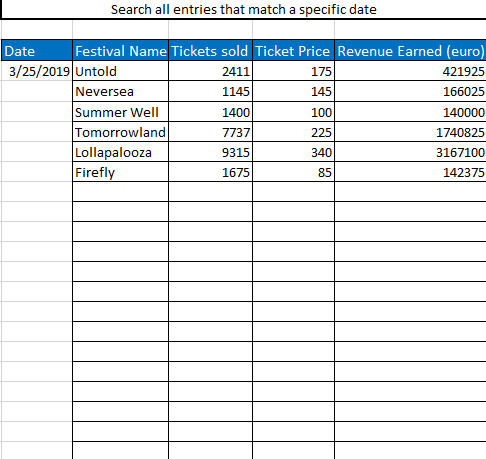


Fig 2.1 Search table output

Conclusion: On March 25th 2019 only six festivals registered sales of tickets.

The problem: Based on the previous table the company must compute the taxes from revenue that occur from ticket sales.

The solution: Using the VLOOKUP function to search the festival name and apply the tax rate to find the tax revenue.



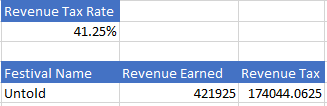


Fig 2.2 VLOOKUP table output

**3. Pivot table**

The problem: The company wants to analyse the revenue earned by each location of the festival.

The solution: To solve the problem, we are going to create a pivot table that will have on a column the location of the festivals and it will display the festival and the revenue earned of that festival.

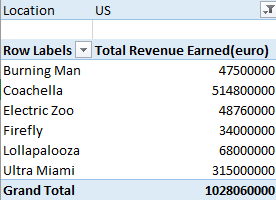


Fig 3.1 Revenue earned from US

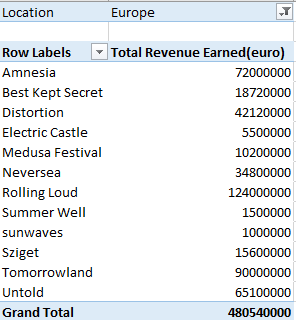


Fig 3.2 Revenue earned from Europe

**4. Creating graphics**

The problem: For the festivals organised by our company, a graphic representing the total ticket sales for each location is being realised. The graphic is constructed based on the pivot table mentioned earlier.

The solution: Based on the previous work, we will generate a Pie chart that will show in percentages the total tickets sales for each festival from a specific location.

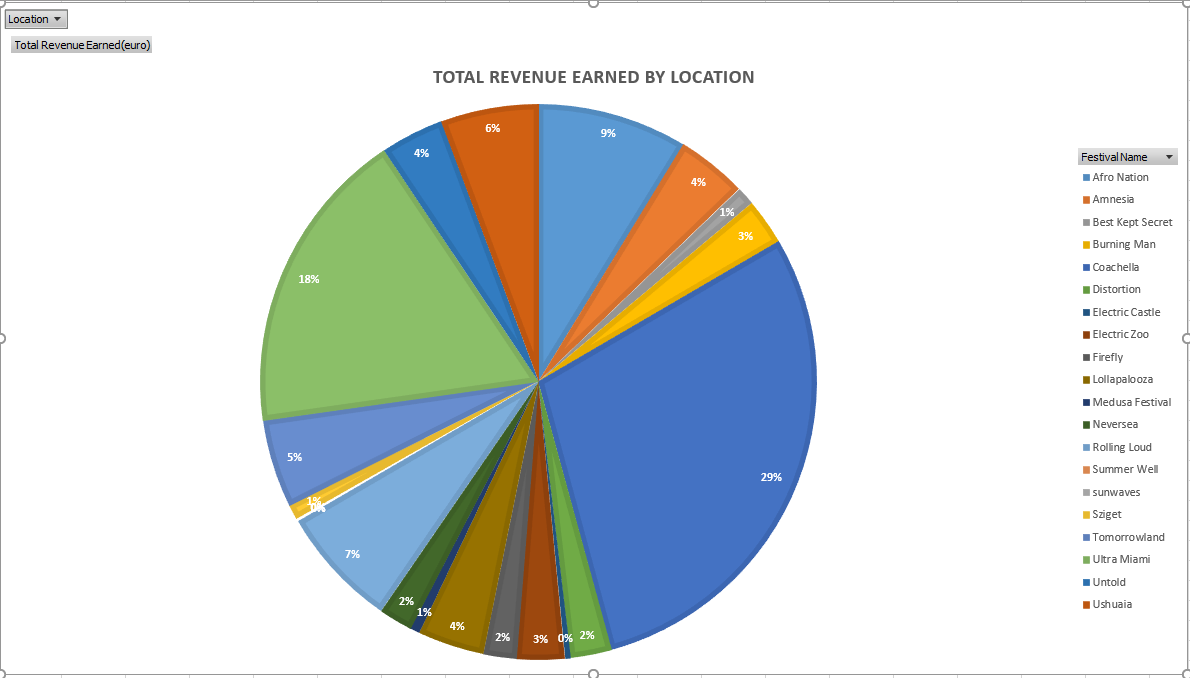


Fig 4.1 Revenue earned from all locations

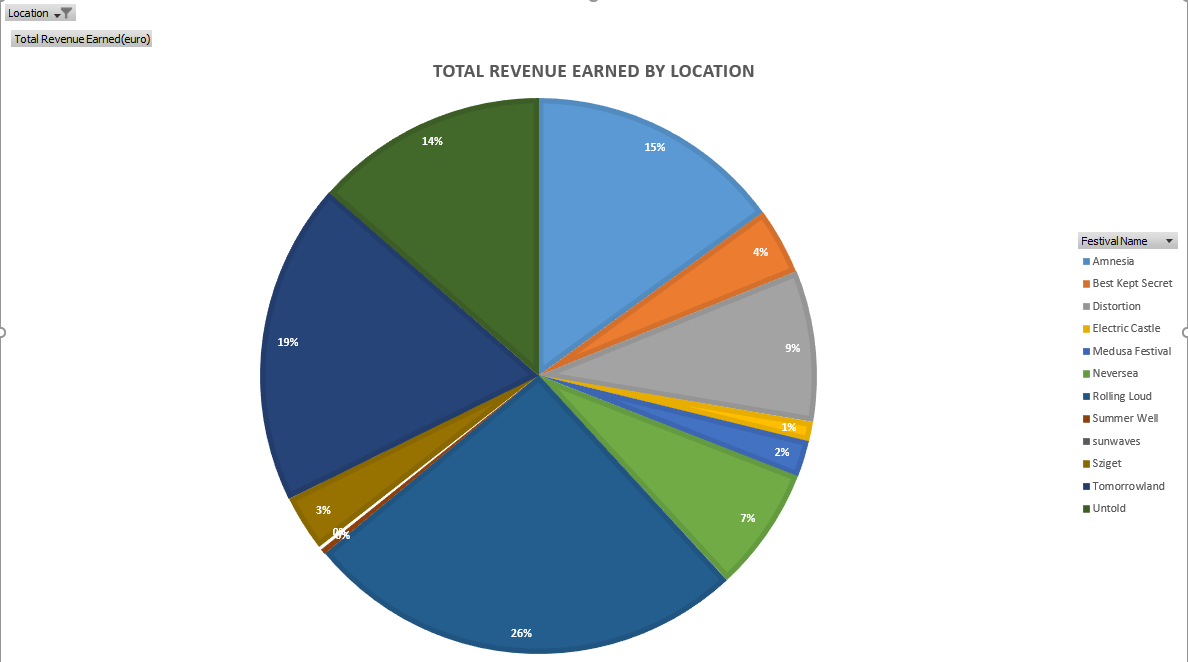


Fig 4.2 Revenue earned from Europe

**Chapter 2.1 SAS Programming**

1. Creating SAS datasets from external files + conditional processing of data

The problem: The company receives data from an external software regarding the festivals to be organised. Those festivals are received in a txt file and are structured as follows: FEST\_ID, FEST\_NAME, FEST\_LOC (EU - Europe, US - United States, OTH - other location), TICKET\_PRICE. Each festival is represented on a new line, and their values are separated by comma.

First, we want to store the data from the txt file in a permanent table. In the same time, we want to change the add a new column called ‘TICKET\_PRICE\_CATEGORY’ that will store the values: ‘NORMAL’ if the ticket price is lower than 100, ‘EXPENSIVE’ if the ticket price is between 100 and 250, and ‘HIGH-END’ if the value of the ticket is greater than 250.

The solution: We read the data from the file and store it in a permanent table, using at the same time the IF/ELSE  IF conditions to verify the price value and to determine the value that will be added in the ‘TICKET\_PRICE\_CATEGORY’ column. And the results are the following:



Fig 2.1.1.1 Creating a dataset

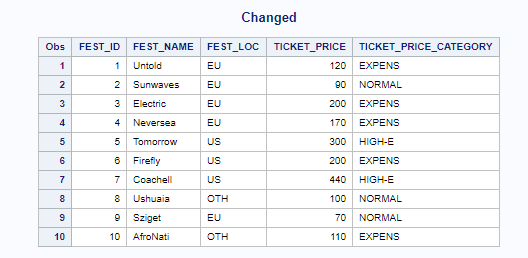


Fig 2.1.1.2 Changing the existing dataset

1. Creating user defined-formats

The problem: When we display the festivals we want to change the text from the column ‘FEST\_LOC’ because it can be misleading. In that way we want to decode the values and show ‘Europe’ for ‘EU’, ‘United States’ for ‘US’ and ‘Other Locations’ for ‘OTH’.

The solution: We define a user-defined format for the column ‘FEST\_LOC’ and we obtain the following results:

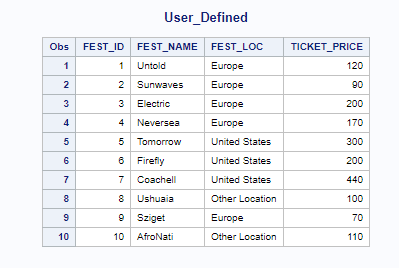


Fig 2.1.2.1 User-defined formats

1. Creating graphics

The problem: The company wants to see a graphical representation of the ticket prices for each festival

The solution: We generate a Graphic from the Graph section in ‘Tasks and Utilities’ section.

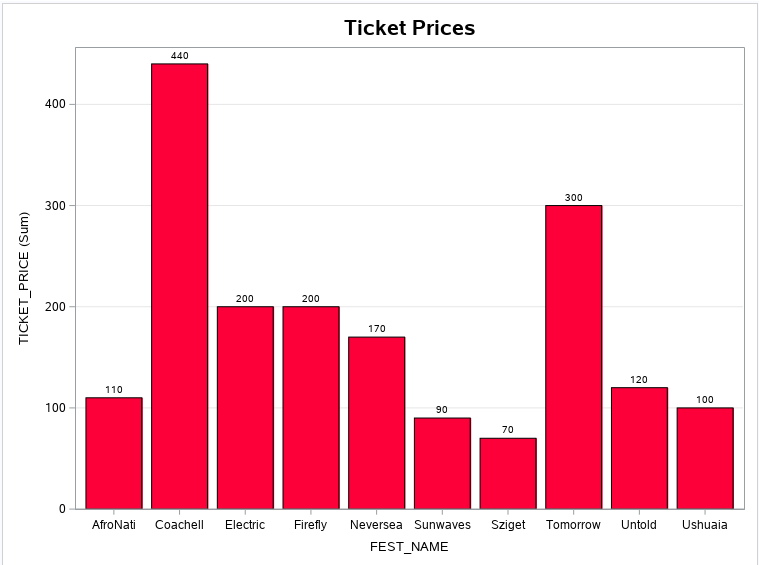


Fig 2.1.3.1 Graphic output

1. Statistics procedures

The problem:The company wants a statistical summary of the ticket prices.

The solution: We generate a statistical summary from the Statistics section ‘Tasks and Utilities’.

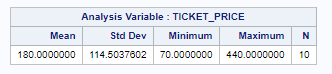


Fig 2.1.4.1 Statistical Summary

1. SQL procedures

The problem:     From the current dataset, query the festivals that will take part in Europe

The solution: Create an SQL procedure that will query the festivals from Europe. And the output will be the following:



Fig 2.1.5.1 SQL procedure output

**Chapter 2.2 SAS Enterprise Guide**

1. The problem: Importing Excel to SAS EG.

The solution: After creating the Process Flow in a new project, right click, add an existing item, and import the excel file. Now right click on the excel icon and select ‘Import data’ and follow the steps.

1. Queries

The problem: Display the days in which the ticket sales revenue was higher than $250.000

The solution: From the imported data, right click, go on ‘Query Builder’ and provide the necessary columns and on the filter tab provide the conditions for the revenue.

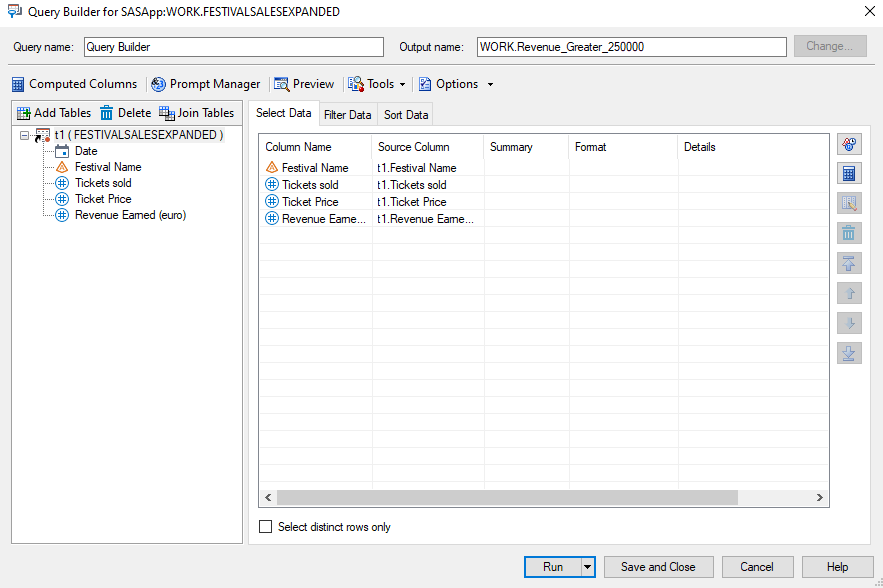


Fig 2.2.2.1 Query builder

1. Reports

The problem: From the previous table create a report.

The solution: Right click on the previous generated table, select ‘Browse tasks’ and search for ‘List Report Wizard’ and follow the steps.

1. Charts

The problem: From the previous table create a chart

The solution: Right click on the previous generated table, select ‘Browse tasks’ and search for ‘Pie Chart’ and follow the steps.

1. Prompts

The problem: Find all tickets sales for a festival, based on the name entered from a prompt.

The solution: Create a prompt from Prompt Manager, found in Query Builder. Create a prompt called ‘Festivals’. Create a Query Builder with the associated columns and on the filter tab select the ‘Festival Name’ column from where the filter will be done and on select the operator ‘Equal to’ as well as the checking the ‘Generate filter for a prompt value’ and select the prompt created previously.

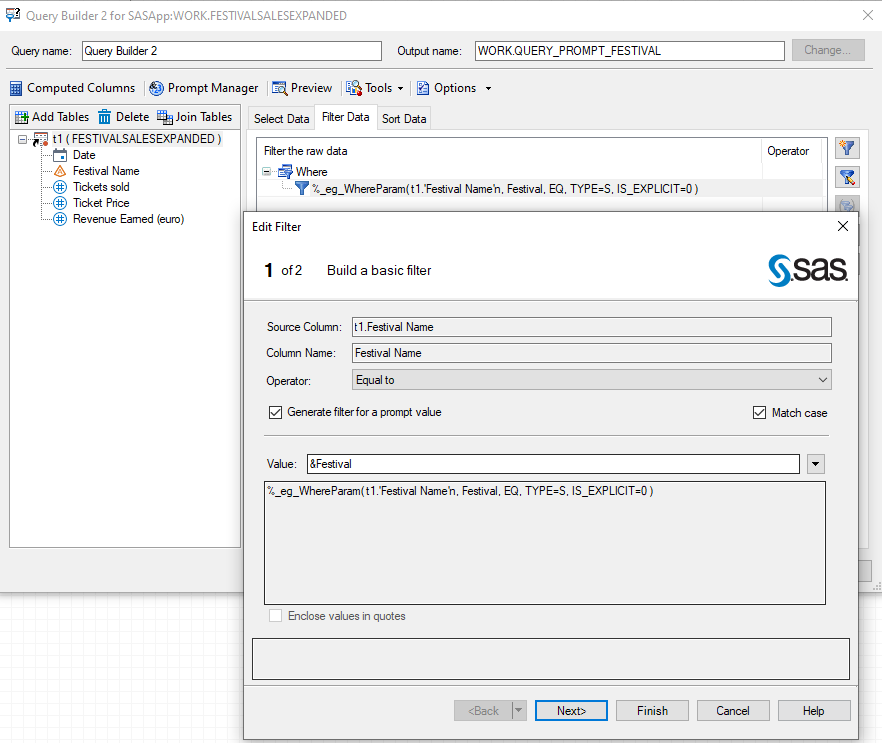


Fig 2.2.5.1 Placing prompts in filter tab

1. Statistical processing

The problem: Create a statistical summary for the previous generated table.

The solution: Right click on the previous generated table, select ‘Browse tasks’ and search for ‘Statistical process’ and select ‘Summary statistics’ and follow the steps.

**Chapter 3. Python**

The scenario in this chapter will assess the trend evolution of the sold tickets over a period of time for each festival that the company is organizing.  The dataset contains a number of 20 festivals where it is recorded the daily number of tickets sold for each festival. The timeline of the analysis is from March 1st 2019 - May 9th 2019 (a total of 70 days).

1. Importing the dataset.

The first problem requires to read the dataset file into a variable to make possible the graphical representations. In this way, to fulfill the request the pandas module comes in handy and reading the file is done using a dedicated function.

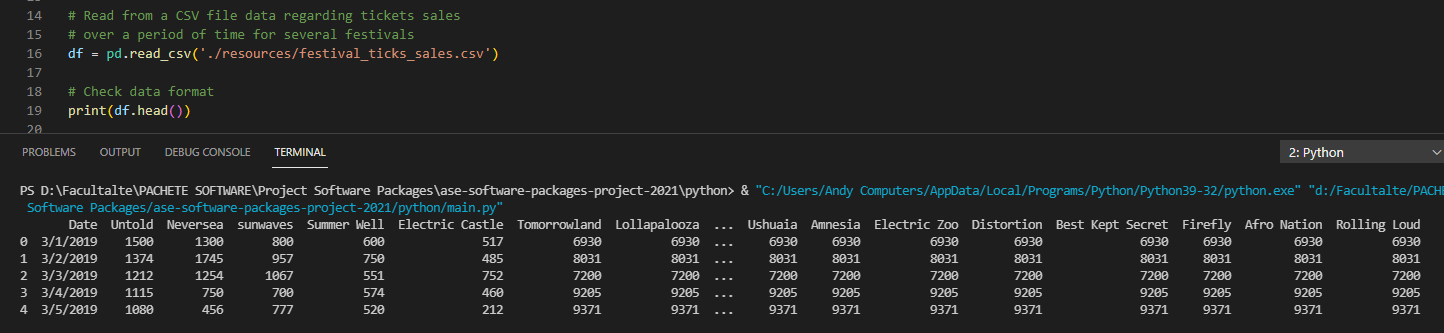


Fig 3.1 Importing the dataset

1. Plot the dataset to assess the evolution of each festival.

The next problem requires to show in a graphical representation the daily number of tickets sold for each festival on the analysed period. To solve the problem it is required to use the matplotlib library in order to have a graphical representation of the data.

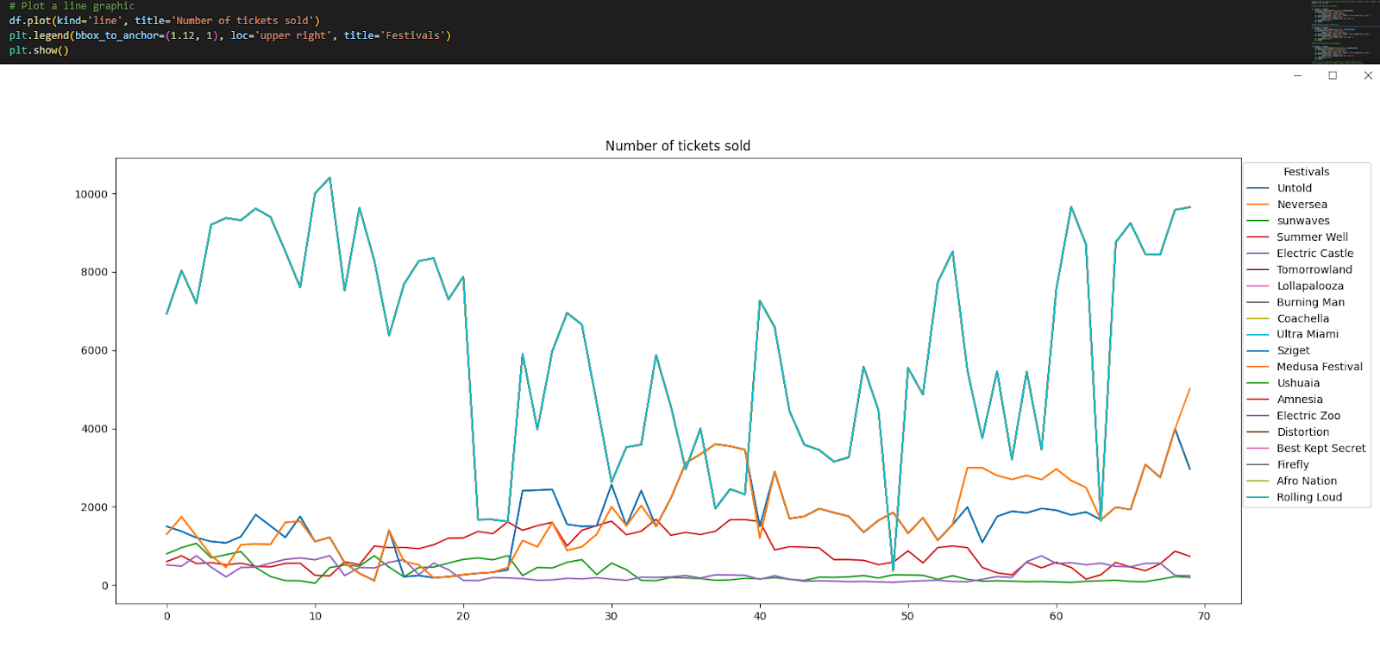


Fig 3.2 Number of tickets sold for each festival

1. Assess the trend evolution of a festival over several time windows.

The following problem involves representing the trend evolution of a festival on different time intervals. To solve the problem, it will be required to use the matplotlib library and also to define a function that returns rolling standard deviation and rolling mean.

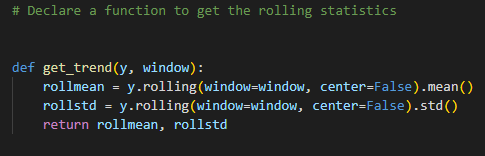


Fig 3.3.1 Function that returns the rolling statistics

After that, using the values from the function the graphics can be plotted.

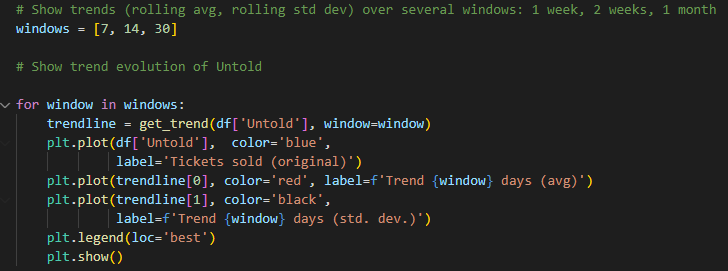


Fig 3.3.2 Plotting the trends over several time windows

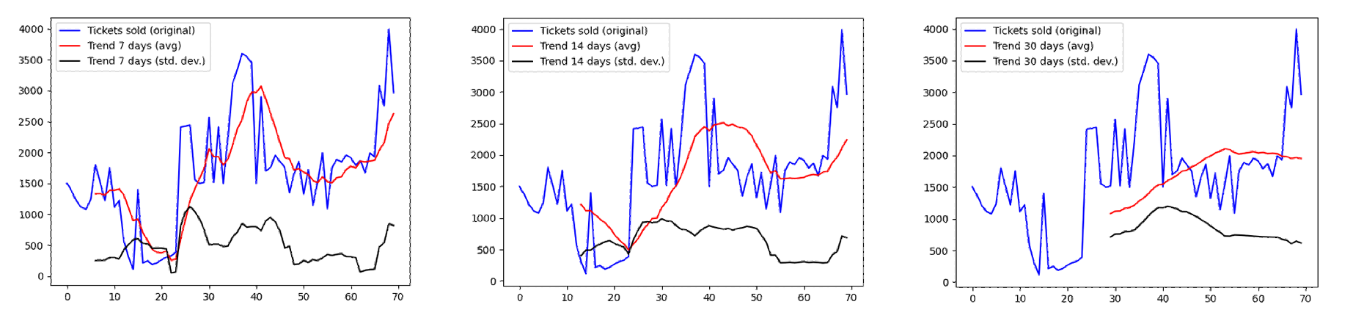


Fig 3.3.2 Trend evolution over one week, two weeks, one month of Untold

1. Check if the number of tickets sold for the Untold festival are related to the number of tickets sold for the Tomorrowland, Neversea and sunwaves festivals.

Solving this problem requires creating a regression model to verify if the Untold variable depends on the variable Tomorrowland,Neversea and sunwaves. And the problem will be solved using the statsmodels library using the ols() function.

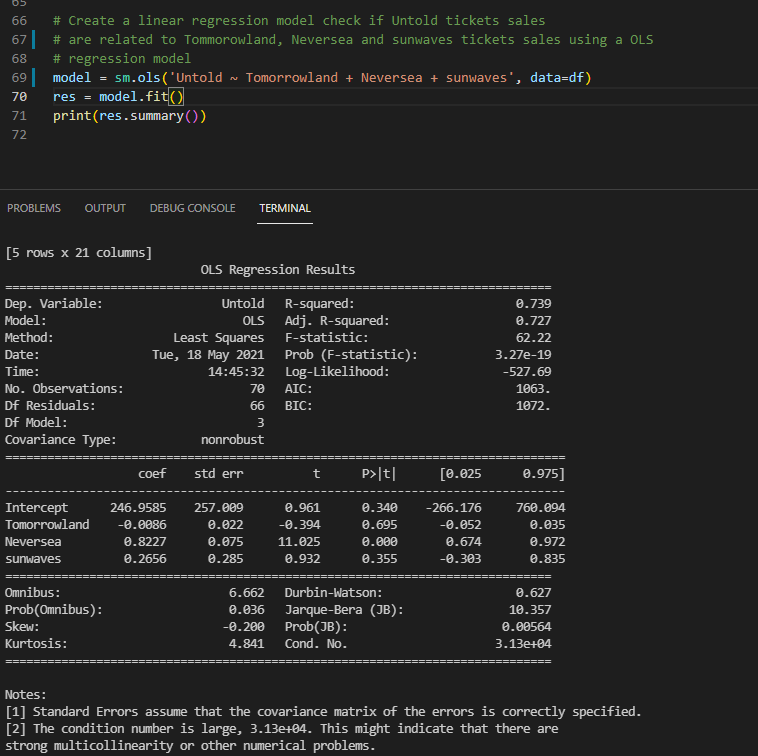


Fig 3.4 Output of the regression model