TABLE OF CONTENTS

TOPICS:

1. Introduction………………………………………………………………
2. Scope of data analysis……………………………………………………
3. Existing System…………………………………………………………
4. Source of dataset………………………………………………………….
5. ETL process…………….…………………………………………………
6. Analysis on database………………………………………………………
7. List of Analysis with Result………………………………………………
8. References…………………………………………………………………..

**OBJECTIVE :**

1. **PRODUCT WISE ANALYSIS**
2. **QUANTITY WISE ANALYSIS**
3. **REGION WISE ANALYSIS**
4. **SHIPPING DETAILS**
5. **COUNTRY WISE SALES**

1.**INTRODUCTION:**

**Ecommerce:**

Ecommerce, also known as electronic commerce or internet commerce, refers to the buying and selling of goods or services using the internet, and the transfer of money and data to execute these transactions. Ecommerce is often used to refer to the sale of physical products online, but it can also describe any kind of commercial transaction that is facilitated through the internet.

Whereas e-business refers to all aspects of operating an online business, ecommerce refers specifically to the transaction of goods and services.

The history of ecommerce begins with the first ever online sale: on the August 11, 1994 a man sold a CD by the band Sting to his friend through his website Net Market, an American retail platform. This is the first example of a consumer purchasing a product from a business through the World Wide Web—or “ecommerce” as we commonly know it today.

Since then, ecommerce has evolved to make products easier to discover and purchase through online retailers and marketplaces.  Independent freelancers, small businesses, and large corporations have all benefited from ecommerce, which enables them to sell their goods and services at a scale that was not possible with traditional offline retail.

[**BUSINESS INTELLIGENCE:**](https://towardsdatascience.com/data-science-for-startups-business-intelligence-f4a2ba728e75)

Identifies common practices for ETLs, automated reports/dashboards and calculating run-the-business metrics. Presents an example with Database Managements and Data Studio(graphs).

**PLATFORM USED IN DATA ANALYSIS (DATA MANAGEMENT):**

1. Microsoft Excel 2016

**VISUALIZE:**

a). Pivot Table

b)**.** Simple graphs and Pictorial representations.

**Business Scenario**

The Analytics team of an Online E-Commerce Company wants to design a Sales dashboard to analyse the sales based on various product categories. The company wants to add user control for product category, so users can select a category and can see the trend month-wise and product-wise accordingly.

The Analytics team also wants to create a histogram to analyse number of shipping days.

The company’s database keeps track of the following data fields:

Order Id Order Date Ship Date Aging Ship Mode Product Category Product Sales Quantity Discount Profit Shipping Cost Order Priority Customer Id Customer Name Segment City State Country Region Months

2.**SCOPE OF DATA ANALAYSIS:**

Data analytics is a process through which data is cleaned, analysed and modelled using tools. This data is then used to derive insights. The insights are then used for business related decision-making purposes. There are many techniques that data analysts use in different fields of work. In the world of business, Data analytics is used for making strategies to get the desired business results. Today, data analytics has become a big career option in India. As a result, big data analytics courses are in huge demand.

Businesses have realised the importance of utilising big data analytics to maximize their profits. They know that it is vital for their growth and for the future health of their business. Today, major business decisions are taken by utilising the insights derived from data related to the organization or industry related data. As competition increases and customers are flooded with choices, it has become important to move faster in the market and that too with accuracy.

**3.EXISTING DATA SYSTEM**

Once you have collected and analysed some of the data, you need to begin to think about how to show it. The first thing you should try to do, is to focus on the story you want to tell. There are a few options, was a prejudice against lower class passengers, or men? Another angle would be to find your odds of surviving if you were a passenger on the Titanic. We also have other data that wasn’t analysed such as ticket price. Did a better ticket buy you a better chance of surviving?

All of these are valid and potential starts for displaying the data. But the main point is that you need to pick one! Throwing a massive data dump at your readers’ feet won’t actually help them to understand anything, and we’re all guilt of doing it. How many times do we get so excited about the dataset we want to show it off in its entirely. Trust me, you don’t have to do that! The data that was used to get our previous numbers also contains the lifeboat number in which the survivors were rescued. We could have used that data, but sometimes it is more important to leave something out, than to try and work it in.

**It should be flexible and scalable**

As we mentioned earlier, no two travel businesses are the same. Every business has individual needs, size and way of doing business. In addition, no single software will satisfy all requirements. Over a period of time as the business and data grows there will need for other software which may need to interoperate with your existing software. Therefore any software that may fulfill your needs today should be flexible and customizable.

Customization is also particularly important for larger companies, which generally do not want an out-of-the-box type of solution.

It should have strong back-office functionality

**It should have strong back office functionality**

One of the major headaches for a small or medium travel agency is the everyday administrative and accounting tasks. They are often complex, time-consuming and counterproductive.

**4.SOURCE OF DATA:**

[ww.kaggle.com](http://www.kaggle.com)

Data drawn from that following website

5**. ETL(Extraction-transformation-loading) PROCESS**

ETL gained popularity in the 1970s when organizations began using multiple data repositories, or databases, to store different types of business information. The need to integrate data that was spread across these databases grew quickly. ETL became the standard method for taking data from disparate sources and transforming it before loading it to a target source, or destination.

In the late 1980s and early 1990s, data warehouses came onto the scene. A distinct type of database, data warehouses provided integrated access to data from multiple systems – mainframe computers, minicomputers, personal computers and spreadsheets. But different departments often chose different ETL tools to use with different data warehouses. Coupled with mergers and acquisitions, many organizations wound up with several different ETL solutions that were not integrated.

Over time, the number of data formats, sources and systems has expanded tremendously. Extract, transform, load is now just one of several methods organizations use to collect, import and process data. ETL and ELT are both important parts of an organization’s broader data integration strategy.

**Extraction:**

During extraction, the desired data is identified and extracted from many different sources, including database systems and applications. Very often, it is not possible to identify the specific subset of interest, therefore more data than necessarily must be extracted, so the identification of the relevant data will be done at a later point in time. Depending on the source system's capabilities (for example, operating system resources), some transformations may take place during this extraction process. The size of the extracted data varies from hundreds of kilobytes up to gigabytes, depending on the source system and the business situation. The same is true for the time delta between two (logically) identical extractions: the time span may vary between days/hours and minutes to near real-time. Web server log files, for example, can easily grow to hundreds of megabytes in a very short period.

**TRANSFORMATION:**

The second step in any ETL scenario is data transformation. The transformation step tends to make some cleaning and con- forming on the incoming data to gain accurate data which is correct, complete, consistent and unambiguous. This process includes data cleaning, transformation, and integration. It de- ﬁnes the granularity of fact tables, the dimension tables, DW(data warehouse) schema , derived facts, slowly changing dimensions, fact less fact tables. All transformation rules and the resulting schemas are described in the metadata repository.

This step include cleaning, filtering and validating and applying rules to extracted data. The main objective of this steps is to load the extracted data into target database with clean and general format

**The process of transformation**: -

Cleaning :-In cleaning the process the dataset is assign with different short cut Such as embarked port where the port is identities by the format like-- “c ” ,“s”,, the data is clean for further demand where it needed .This shows that name of the port where the ship has to embarked, whereas the male is assign with “M” and the female is assign with the “F”.

Filtering: - In filtering, dataset is selected, or we say that attributes are select from the source csv file and to analysis the taken attributes to get the information in the graphical or in the tabular new data, this also shows that how the data is converted into shortening and the new tabular gives focused and vast information for that attributes.

Splitting: -In splitting, the source data (in csv format) has used the pivot table in which multiple column has taken to show the data in the pivot table and after that pivot table has converted in graphical form to analyse data.

Enrichment: -Here the source data has certain attributes like the data which we have taken has “first name” and “last name” these notation gives enrichment to the data. Where the information is gathered from the notation, it is the best way to write the name of the person in place of writing “full name” we can enrich into “first name” and the “last name”.

**LOADING:**

This is the final step in the ETL process. In this step, the extracted data and transformed data is loaded to the target database. In order to make data load efficient, it is necessary to index the database and disable constraints before loading the data.

All the three steps in the ETL process can be run parallel. Data extraction takes time and so the second step of transformation process is executed simultaneously. This prepares data for the third step of loading. As soon as some data is ready it is loaded without waiting for completion of the previous steps.

6**. Analysis on database:**

**Overview**

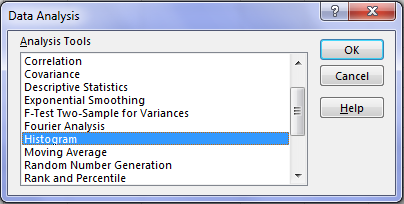
* Use the Saved Sample – E-Commerce database.
* Prepare a table of Sales and Profit month-wise in working sheet.
* Prepare the sales table region-wise in the working sheet.
* Create User Control Combo box for Product Category.
* Create Column Chart of month-wise table and region-wise table.
* Link the table with combo box.
* Create a dashboard.

**Steps:**

**Step1: Create Histogram for Shipping Days(Aging)**

To create histogram, click the Data Tab, Under Analysis Group (Right Corner), Click Data Analysis.

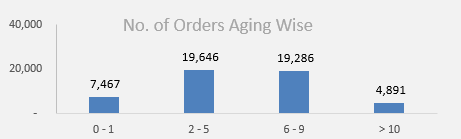
Now, select Histogram and click ok. A histogram dialog box will appear.



In the histogram dialog box, first click the Label’s Check box as we have labels in our data. After that, In the **Input reference box** select the range **(“Sales Data!D1:D51291”)** of our dataand in the **Bin Range Reference box** select **(“Working!K3:K7”)**.

In **Output section**, select range “Working!N3” for binning table, click Histogram check box and then ok.



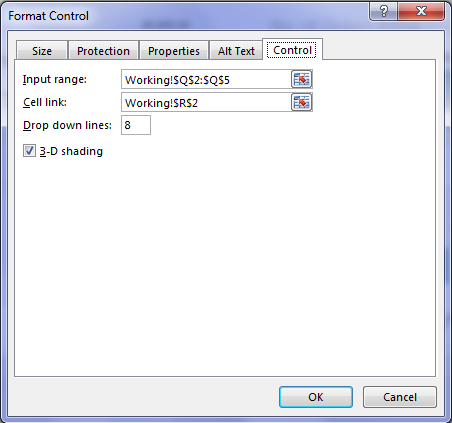
****

**Step2 - Create Combo Box:**

* Insert Combo box for product category list in the Dashboard Sheet.
* Click Developer Tab > Under Controls Panel > Click Combo box and draw.

Pass the Input Range and Cell for the Combo box.

* Right-click the country list Combo box > Click Format Control > Under Format Control Panel, Pass Input Range “Working!Q2:Q5” and Cell Link “Working!R2” from the working sheet.

****

Now, write the offset function in cell “R3” to fetch the product category based on the selection in the product category Combo box.

* Write the equal sign and then the function name.
* Pass the first argument as Cell “$Q$1.”
* In the second argument, select the cell link cell “$R$2.”

**Step3: SUMIFS formula to calculate Total Sales, Quantity, and Profit**

Now, write Sumifs formula to calculate Sales, Quantity, and Profit in the Dashboard sheet.

Enter the formula in Cell C7:

* Enter the equal sign and then enter the function name and open parenthesis.
* Pass the first Argument is Sum\_Range, select range ‘Sales Data’!$H:$H, and then enter comma.
* Now, pass the second argument Product Category column “criteria Range1” as ‘Sales Data’!$F:$F, enter comma
* Pass the third argument “criteria1” “Working!$R$3”, and enter comma.

****

**Perform the same function to calculate the Quantity in Cell G7.**

* In G7, write the equal sign, and then enter the function name and open parenthesis.
* The first Argument is Sum\_Range, select range ‘Sales Data’!I:I, and then enter comma.
* Now, pass the second argument Product Category column “criteria Range1” as ‘Sales Data’!F:F, and enter comma.
* Pass the third argument “criteria1” “$R$3,” and enter comma.

**For Profit**

* In K7, write the equal sign and then enter the function name and open parenthesis.
* The first Argument is Sum\_Range, select range ‘Sales Data’!K:K, and then enter comma.
* Pass the second argument Product Category column “criteria Range1” as ‘Sales Data’!F:F, and enter comma.
* Now, pass the third argument “criteria1” “$R$3”, and enter comma.

**Step4: SUMIFS formula to calculate Sales and Profit month wise**

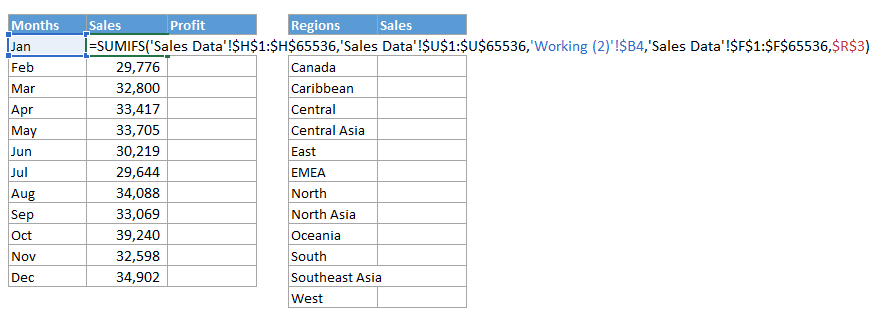
Now write the sumifs formula to calculate the Sales and profit month-wise and sales region-wise.

Enter formula in Cell C4:

* Enter the equal sign and then enter the function name and open parenthesis.
* The first Argument is Sum\_Range, select range ‘Sales Data’!H:H, and then enter comma.
* Pass the second argument month column “criteria Range1” as ‘Sales Data’!U:U, and enter comma.
* Now, pass the third argument “criteria1” “$B$4,” and enter comma.
* Pass the fourth argument as Data!F:F product category column, and enter comma.
* Pass the fifth argument as “$R$3.”
* Now, copy and paste the formula in Range C4:C15.

Enter formula in Cell D4:

* Enter Equal sign then enters function name and open parenthesis
* The first Argument is Sum\_Range, select range ‘Sales Data’!K:K, and then enter comma.
* Now, pass the second argument month column “criteria Range1” as ‘Sales Data’!U:U, and enter comma.
* Pass the third argument “criteria1” “$B$4,” and enter comma.
* Pass the fourth argument as Data!F:F product category column, and enter comma.
* Enter the fifth argument as “$R$3.”
* Now, copy and paste the formula in Range D4:D15.

****

**Step5: SUMIFS formula to calculate Sales region wise**

* Write the equal sign and then enter the function name and open parenthesis.
* The first Argument is Sum\_Range, select range ‘Sales Data’!H:H, and then enter comma.
* Pass the second argument region column “criteria Range1” as ‘Sales Data’!T:T, and enter comma.
* Now, pass the third argument “criteria1” “$F$4,”and enter comma.
* Pass, the fourth argument as Data!F:F product category column, and enter comma.
* Pass the fifth argument as “$R$3.”
* Now, copy and paste the formula in Range G4:G15.

**Step 6: Create Column Chart**

Now, create the column chart for both region-wise and month-wise table.

Select table (B3:D15), click insert tab > under Charts Panel > Insert column chart.

Cut and Paste the chart in the Dashboard Sheet.

Perform the same steps for other tables to create chart.

Now, this is our sales Dashboard, we can apply any color in the interior of cells, and data series to format it.

**7.List of Analysis with Result:**

**SAMPLE FORM OF DASHBOARD**

****

Sales and profits analysis monthly wise

|  |  |  |
| --- | --- | --- |
| **Months** | **Sales** | **Profit** |
| Jan | 87,526 | 38,447 |
| Feb | 85,683 | 38,753 |
| Mar | 95,249 | 41,165 |
| Apr | 95,962 | 42,366 |
| May | 91,445 | 40,872 |
| Jun | 96,597 | 42,010 |
| Jul | 91,690 | 40,959 |
| Aug | 88,153 | 38,238 |
| Sep | 89,216 | 39,656 |
| Oct | 97,347 | 43,305 |
| Nov | 88,575 | 39,016 |
| Dec | 89,696 | 39,490 |

Monthly wise quantity analysis

|  |  |
| --- | --- |
|  | QUANTITY |
| **Jan** | 12985 |
| **Feb** | 11501 |
| **Mar** | 13349 |
| **Apr** | 12314 |
| **May** | 12854 |
| **Jun** | 12851 |
| **Jul** | 13277 |
| **Aug** | 13122 |
| **Sep** | 12568 |
| **Oct** | 13087 |
| **Nov** | 12683 |
| **Dec** | 13141 |

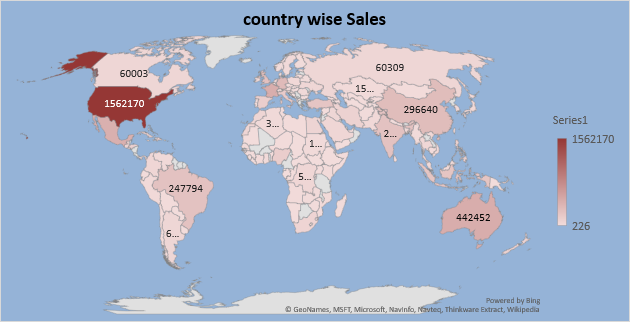
Region wise Sales and Profits analysis

Shipping Days (Aging) Analysis

Sales and Profits analysis Product wise

|  |  |  |
| --- | --- | --- |
| **Product Name** | **Profit** | **Sales** |
| Apple Laptop | 33025 | 55250 |
| Bed Sheets | 114958.49 | 216908 |
| Beds | 35752.6 | 80262 |
| Bike Tyres | 26803.2 | 59472 |
| Car & Bike Care | 22700.84 | 97468 |
| Car Body Covers | 21618.52 | 96642 |
| Car Mat | 20836.8 | 44604 |
| Car Media Players | 39989.6 | 115640 |
| Car Pillow & Neck Rest | 107872.48 | 191499 |
| Car Seat Covers | 20006.9 | 94278 |
| Car Speakers | 92146.79 | 174286 |
| Casula Shoes | 86332.46 | 340990 |
| Curtains | 16101.83333 | 34952 |
| Dinner Crockery | 43340.25 | 140049 |
| Dinning Tables | 28867.65 | 122451 |
| Fans | 12045.5 | 32929 |
| Formal Shoes | 318363.59 | 595335 |
| Fossil Watch | 181680.96 | 444564 |
| Iron | 26833.22 | 49062 |
| Jeans | 331900.4 | 609092 |
| Keyboard | 3290.1 | 7293 |
| LCD | 6193.416667 | 14560 |
| LED | 20910.08 | 43008 |
| Mixer/Juicer | 5030.676667 | 18592 |
| Mouse | 4629.89 | 24531 |
| Running Shoes | 346338.88 | 626080 |
| Samsung Mobile | 26568.6 | 48620 |
| Shirts | 275252.96 | 547624 |
| Shoe Rack | 34018.04 | 127596 |
| Sneakers | 80057.5 | 173290 |
| Sofa Covers | 119393.92 | 221832 |
| Sofas | 30738.48333 | 68943 |
| Speakers | 9909.2 | 33930 |
| Sports Wear | 47239.45 | 240295 |
| Suits | 53725.43 | 304655 |
| T - Shirts | 407716.88 | 692912 |
| Tablet | 22311.04 | 43979 |
| Titak watch | 355244.12 | 637260 |
| Towels | 131158.88 | 234384 |
| Tyre | 132302.5 | 223250 |
| Umbrellas | 33267.5 | 72030 |
| Watch | 3428.32 | 22984 |

Country wise Sales



8**.References:**

MICROSOFT EXCEL 2016 BIBLE:THE COMPREHRNSIVE TUTORIAL RESOURCE (Text Book)

FUNDAMENTALS OF BUSINESS ANALYTICS (Text Book)

www. Easyexcel.com