Power bi mandatory project notes:

To extract, clean, load, and transform the downloaded Excel file, follow these steps:

1. **Load and transform the Excel file:** Open the data extraction tool (e.g., Power Query in Microsoft Excel) and select "Get Data." Choose the option to load the data as a table and click "Transform" to open the data transformation editor.

2. **Splitting the Address column**: In the data transformation editor, locate the Address column and click on "Split Column." Choose the appropriate delimiter (e.g., comma, space and if the required delimiter is not present click on custom and provide the delimiter e.g., dash space) to split the column into separate components such as City, State, Country, and Postal Code.

3. **Checking unique values in the shipping mode column**: Select the Shipping Mode column and right-click on it. Choose the option to view the unique values within the column. This will display a list of all distinct values present.

4. **Replacing values in the shipping mode column**: After identifying the unique values in the Shipping Mode column, click on the "Replace" option from the data transformation toolbar. In the pop-up box, enter the value that needs to be replaced (e.g., "FC") and specify the replacement value (e.g., "First Class"). Confirm the replacement to update the column with the desired values.

Data Modelling:

To transform the data into a star schema for improved analysis performance, the following steps were taken:

1. Converting the data into fact and dimension tables:

a. In the Power BI Power Query Editor, the Order table was duplicated by right-clicking on it (duplicate the table three times for three dimension tables).

b. Each duplicated table was renamed, and only the necessary columns were selected to create the respective dimension tables (e.g., Customer table includes CustomerID, Customer Name, Customer Segment). Unwanted columns were removed by right-clicking and selecting "Remove Other Columns."

2**. Removing duplicate rows from dimension tables:**

Duplicate rows were eliminated from the selected columns to ensure the tables were free from duplicates. For example:

- Duplicate rows in the Customer table were removed by selecting the "Remove Rows" option from the Home tab and then choosing "Remove Duplicate Rows."

- Similar steps were followed for removing duplicate rows based on Postal Code and Product ID in the Order Details and Product tables.

3**. Establishing one-to-many relationships between dimensions and the fact table:**

Once the tables were created, it was important to ensure that the necessary relationships were established between the dimensions and the fact table. Here's how it was done:

- In the model view of Power BI, the relationships between the tables were checked by hovering over the relationship lines to verify if they were correctly established.

- If any relationship was missing or incorrect, it was deleted by right-clicking on the relationship and manually establishing it.

- For example, a one-to-many relationship was created by selecting the Postal Code column from the Order Details table and dropping it into the Order table to establish the connection.

Data Analysis:

To create a new column 'Sales' or 'Order value' and discounted sales in Power BI and display the total sales using a card visual, you can follow these steps:

a. Open your Power BI report and go to the "Fields" pane on the right side.

b. Locate the Order table that contains the columns 'Qty', 'price per each', and Discount'.

c. Select "New column" from the table tool menu.

d. In the formula bar at the top, enter the following formula: Sales = [Qty] * [price per qty] * (1 - Discount)

e. Press Enter to apply the formula. Power BI will calculate the values for the new 'Sales' column based on the existing columns.

f. Go to the report page, to add the card visual, from the Visualizations pane on the right side, select the "Card" visual.

g. Drag and drop the 'Sales' column from the "Fields" pane onto the "Values" field well of the card visual.

h. Power BI will automatically sum up the values in the 'Sales' column and display the total sales in the card visual.

i. After that customization of the appearance of the card visual and canvas was done separately by using the formatting options available in the Visualizations pane.

Similarly calculate the Sales from discounted products and display the total sales from discounted products. Formula used for discounted sales is

```
Discounted Sales = CALCULATE(SUM(Orders[Sales]),Orders[Discount]>=0.1)
```

To create a column “Cart Value” that categorizes the order value/sales as Low, medium, high or very high.

a. New column Cart value was created by same method as mentioned above.

b. Formula used to create cart value is

   Cart value = SWITCH(TRUE(), Orders[Sales]>10000,"Very high",Orders[Sales]>35000,"High",Orders[Sales]>1000,"Medium","Low")

c. Switch True is used in placed of nested IF.

d. Go to the report page, to add the pie chart, from the Visualizations pane on the right side, select the "Pie" visual.

e. g. Drag and drop the 'Cart Value' column from the "Fields" pane onto the "Legend" field and ‘Sales’ in the “Value” field.

f. After that customization of the appearance of the pie chart and was done separately by using the formatting options available in the Visualizations pane.

To separately visualize the total sales from the low cart value category in Power BI, you can follow these steps:

1. In Power BI report page, select card visual.

2. In the "Fields" pane on the right side of the Power BI interface, locate and select the Order dataset which contains the sales data.

3. Drag and drop the "Sales" column into the "Field" area of the visualizations pane.

4. Drag and drop the "Cart Value" column into the "Filter on the visual " area of the Filter pane.

5. In the "Filters" pane, find the "Cart Value" field and click on the drop-down arrow to open the filter options.

6. In the filter options, select and set the filter condition to be “Low Cart”. This will filter out all values above 1000, leaving only the low cart value category.

7.After that customization of the appearance of the pie chart and was done separately by using the formatting options available in the Visualizations pane.

To create a new measure, sales using calculate() and sum() function with the two filters, Low cart category and discount>=50% OR calculate sales using sumx() with if() and and()].

a. Open your Power BI report and go to the "Fields" pane on the right side.

b. Create a new table “Measure” which will contains all measures.

c. Select "New measure" from the table tool menu.

d. In the formula bar at the top, enter the following formula: Discounted Sales = CALCULATE(SUM(Orders[Sales]),Orders[Discount]>=0.5,Orders[Cart value]="Low")

e. Press Enter to apply the formula. Power BI will calculate the values for the new Discounted Sales measure.

To find out the number of days it takes to deliver for each shipment type (refer ship mode) so that delivery issues can be looked at on. Creates a column chart that shows the average number of days it takes to deliver for each shipment type.

a. New column Average Delivery Day was created by same method as mentioned above in the order table in the data view.
b. Formula used to create Delivery Day : Delivery Days = DATEDIFF(Orders[Order Date],Orders[Ship Date],DAY)
c. In Power BI report page, select Clustered column chart.
d. Drag and drop the "Delivery Days" and “Ship mode” column into the chart.
e. After that customization of the appearance of the pie chart and was done separately by using the formatting options available in the Visualizations pane.

To create a matrix visualization that displays order date as hierarchy , sales and sales year to date.

a. Select "New measure" from the table tool menu.

b. In the formula bar at the top, enter the following formula: Total Sales = SUM(Orders[Sales])

c. Press Enter to apply the formula. Power BI will calculate the values for the new Total sales measure.

d. Create another measure “YTD” in the same way by using the formula: YTD = TOTALYTD([Total Sales],Datetable[Date]).

e. Choose matrix visualisation table from the chart, create a date hierarchy in the date dimension table.

f. Locate the date table in the Data pane, right click select date hierarchy, after that keep on adding the column by selecting the columns, right clicking and add to hierarchy from date table.

g. Drag and drop “Total Sales”, “YTD”, and “Date Hierarchy” Measure into the Matrix Visualisation.

h. Remove the unwanted column of date hierarchy in the “column” field in the visualisation pane.

To visualize the cumulative sales for each month for all the years to calculate Year on Year Sales Growth. Calculate YoY growth.

a. Select "New measure" from the table tool menu.

b. In the formula bar at the top, enter the following formula: YOY = CALCULATE([Total Sales], DATEADD('Datetable'[Date],-1,YEAR))

c. Press Enter to apply the formula. Power BI will calculate the values for the new YOY measure.

d. Choose matrix visualisation table from the chart.

g. Drag and drop “Total Sales”, “YOY”, and “Date Hierarchy” Measure into the Matrix Visualisation.

h. Remove the unwanted column of date hierarchy in the “column” field in the visualisation pane.

Insights:

1. **Total sales have been increasing continuously from the year 2014 to 2017.**
2. **Category wise we can see that technology have highest number of sales.**
3. **In the category “technology”, “Phones” have the highest sales.**
4. **Consumer Segment customers are contributing more in the sales, followed by Corporate and Home office.**
5. **Maximum product sales occur in the range below than 1000.**
6. **West part of the region is contributing more to the sales, compared to other parts of the region.**
7. **Time takes to ship the standard mode of delivery is almost 5 days, followed by second class 3 days and first class is 2 days.**
8. **Chances of shipping in the same order day is almost none.**
9. **Low cart sales where discount => 50%, contributed around 14.11K of total low cart sales which is 1.28 M.**
10. **Discounted product sales contribute almost one-fourth of the total sales.**