**Analyzing Data with Excel**

*Sales Data Project*

1. **Cleaning Data : Remove blank and duplicate data, Change Case of the text, Fix date formatting errors, Trims Whitespace**
2. **Data Analysis Basics : Filtering and Sorting data, Using some excel function (**IF, IFS, COUNTIF, SUMIF, VLOOKUP, HLOOKUP)
3. **Pivot Tables : obtain usable and presentable insights into data, Using filter and slicer**

**ANALISIS DATA MENGGUNAKAN MICROSOFT EXCEL**

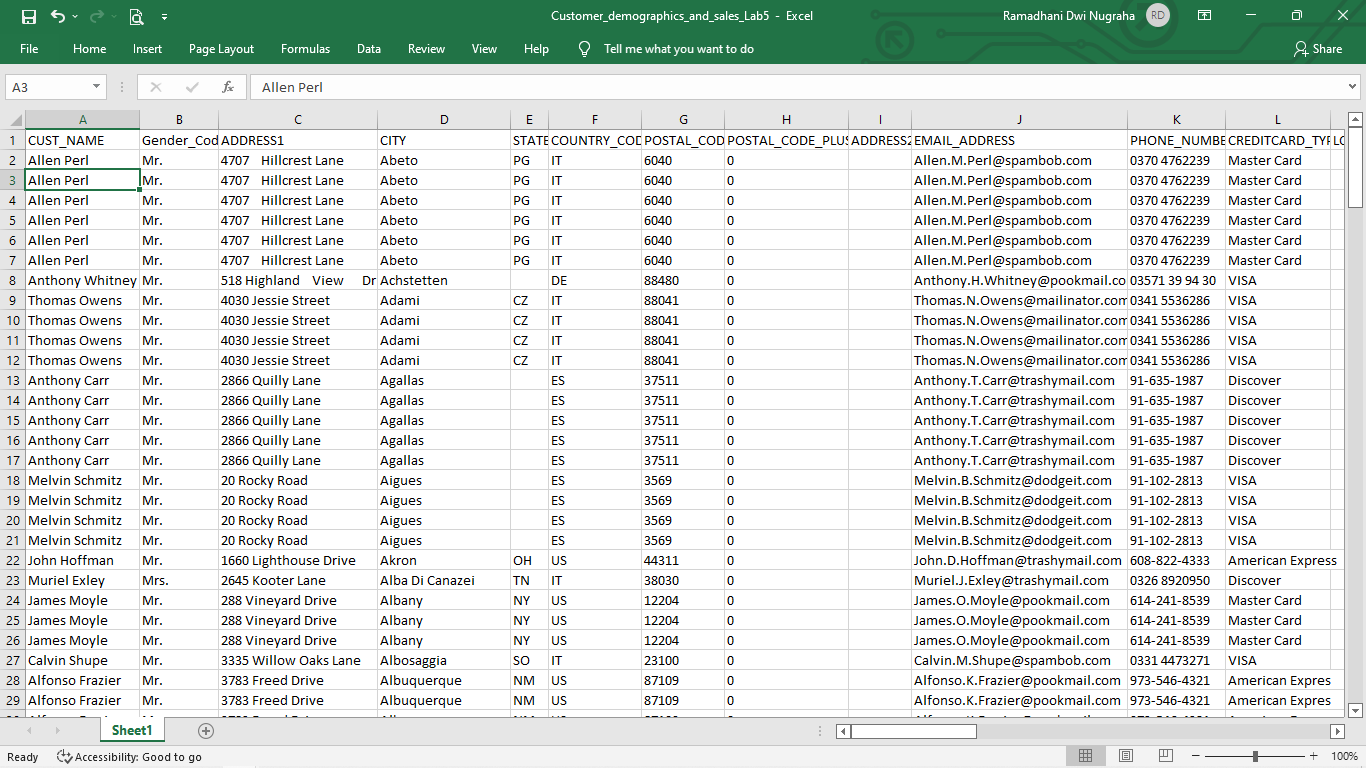
Pengelolaan data dari data mentah sehingga menjadi data yang dapat memberikan *insight* dapat dilakukan dengan bermacam-macam tools. Salah satunya yaitu Microsoft Excel. Pada kesempatan kali ini saya mencoba menganalisis data dari beberapa situs yang menyediakan data secara *open souce*. Adapun outline dari project ini adalah:

* Cleaning Data
* Basic Data Analysis
* Pivot Tables

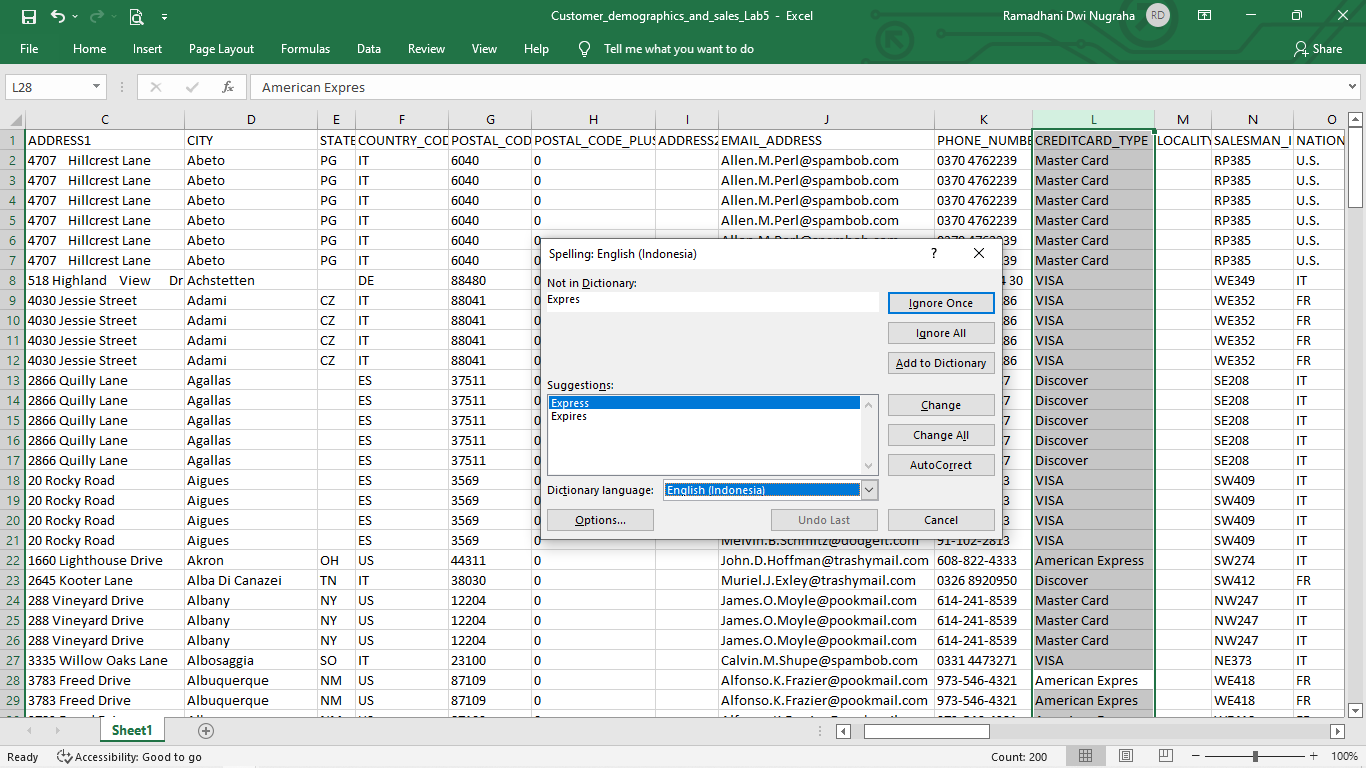
Pada tahapan pertama yaitu *Cleaning Data*. Umumnya, dalam menganalisis sebuah data diperlukan data yang sudah siap pakai untuk memberikan *insight* baru. Namun tidak jarang data yang didapat banyak memiliki duplicated data, inaccurate, ataupun empty rows. Yang dimana kasus-kasus tersebut memiliki dampak terhadap analisis data. Untuk itu, kita akan membahas satu persatu dari kasus yang sering terjadi pada tahapan cleaning data.

1. Irrelevant, Empty, dan Duplicated Data

Data yang digunakan pada kasus ini adalah Customer Demographic and Sales[[1]](#footnote-1)

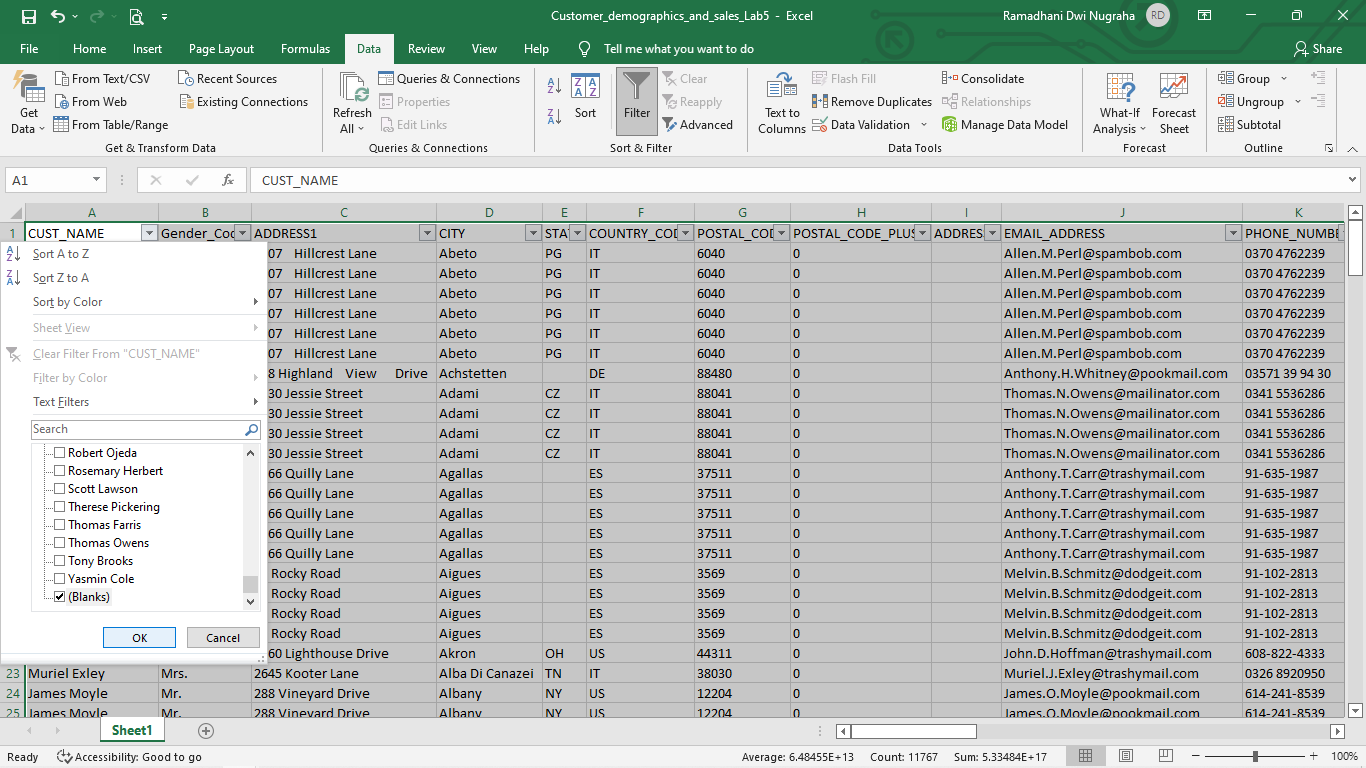


Pada tahapan pertama, kita dapat melakikan pemeriksaan pengejaan pada kolom yang berisi dengan kata-kata umum. Fitur pengejaan ini dapat ditemukan pada tab **Review** lalu **Spelling.** Data yang akan kita coba pengejaannya adalah kolom L yang berisi CREDITCARD\_TYPE.

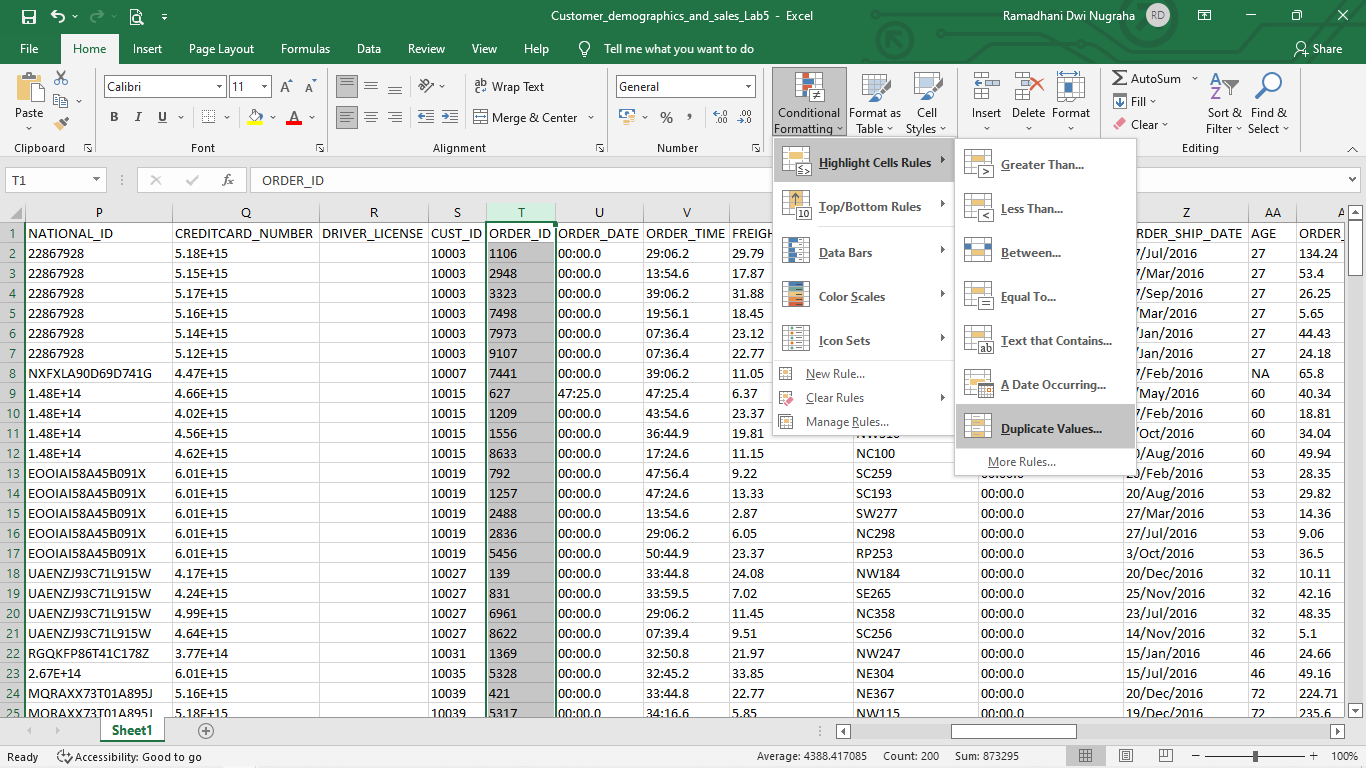


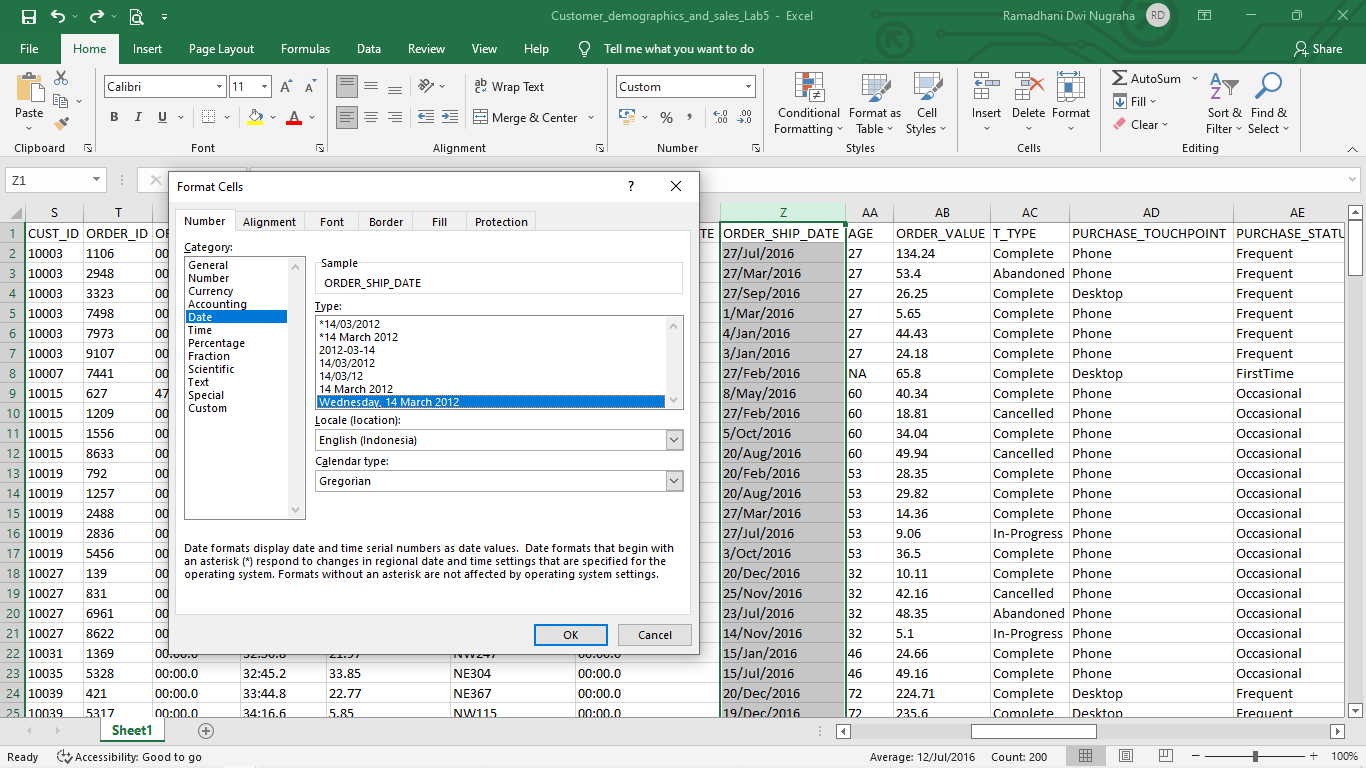
Kolom tersebut memiliki pengejaan data yang tidak tepat pada penulisan Express, sehingga dapat dengan mudah kita ganti dengan memilih Change All.

Kemudian, kita dapat menghapus empty rows yang terdapat pada data excel dengan cara mengaplikasikan Filter pada keseluruhan data dan memilih Blank data untuk menampilkan empyt rows. Fitur filter terdapat pada tab Data, bagian Sort&Filter kita dapat memilih Filter. Untuk lengkapnya dapat dilihat digambar berikut:

Tidak lupa juga, pada saat mengaplikasikan filter pada data kita perlu select keseluruhan data hal ini dapat dilakukan dengan shortcut (CTRL+SHIFT+END). Setelah memilih Blanks pada filter tersebut, kita dapat menghapus rows yang tidak berisi data.

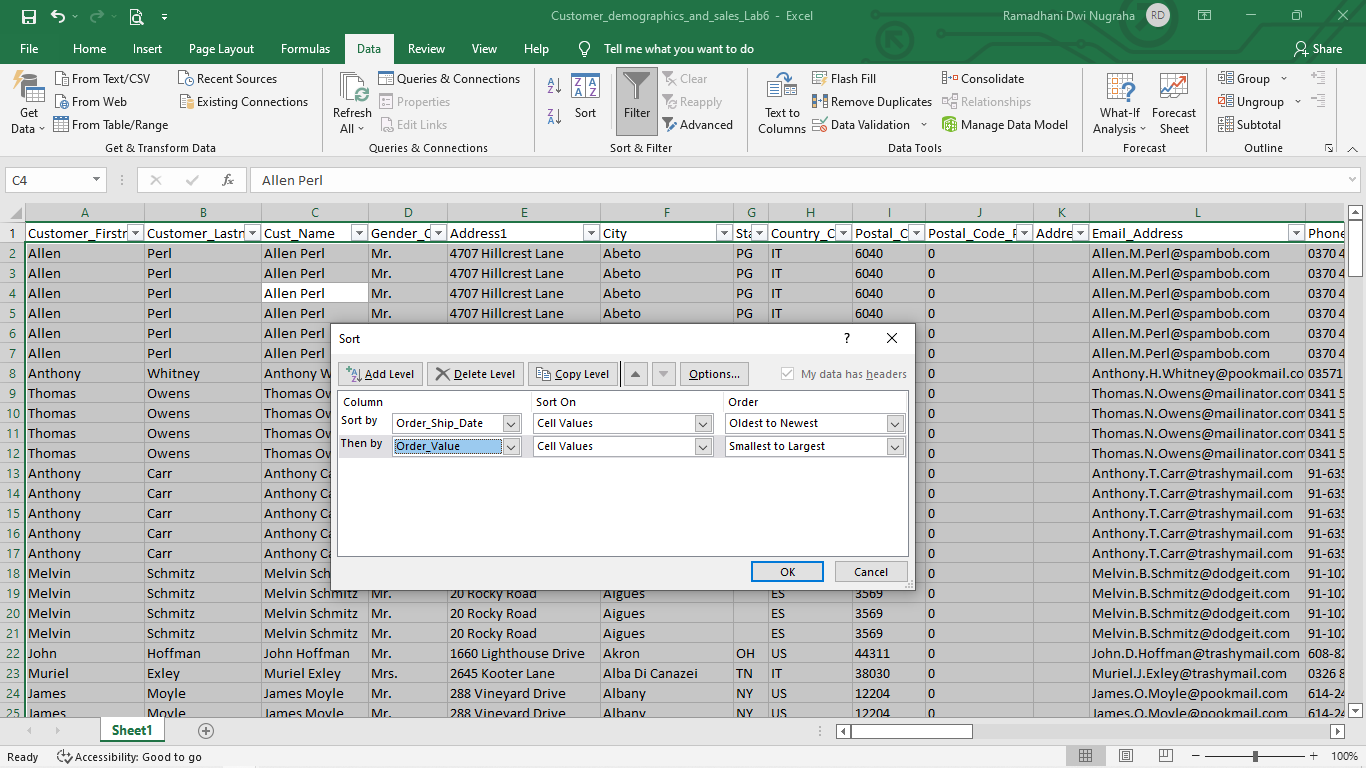
Selanjutnya kita dapat menghapus duplicated data menggunakan fitur **Removes Duplicated** pada tab **Data** bagian **Data Tools.** Untuk melihat data yang memiliki kesamaan isi pada sel lainnya kita dapat menggunakan fitur Conditional Formatting yang menampilkan Duplicate Values. Kita akan menghighlight data yang terduplicate dengan warna yang berbeda dengan lainnya. Data yang akan kita aplikasikan fitur ini adalah data yang terdapat pada kolom T yaitu ORDER\_ID yang dimana data ini merupakan data yang unik dan memiliki perbedaan setiap datanya.

Selain itu, kita juga dapat mengganti format penanggalan yang terdapat pada data menjadi data yang lebih mudah untuk ditangkap informasinya. Hal ini dilakukan pada kolom Z ORDER\_SHIP\_DATE. Fitur teradapt pada tab Home pada bagian Number dan pilih More Number Format dan pilih date. Pada kasus ini dipilih format penanggalan dengan mencantumkan hari pada kolom.



1. Basic Data Analsis

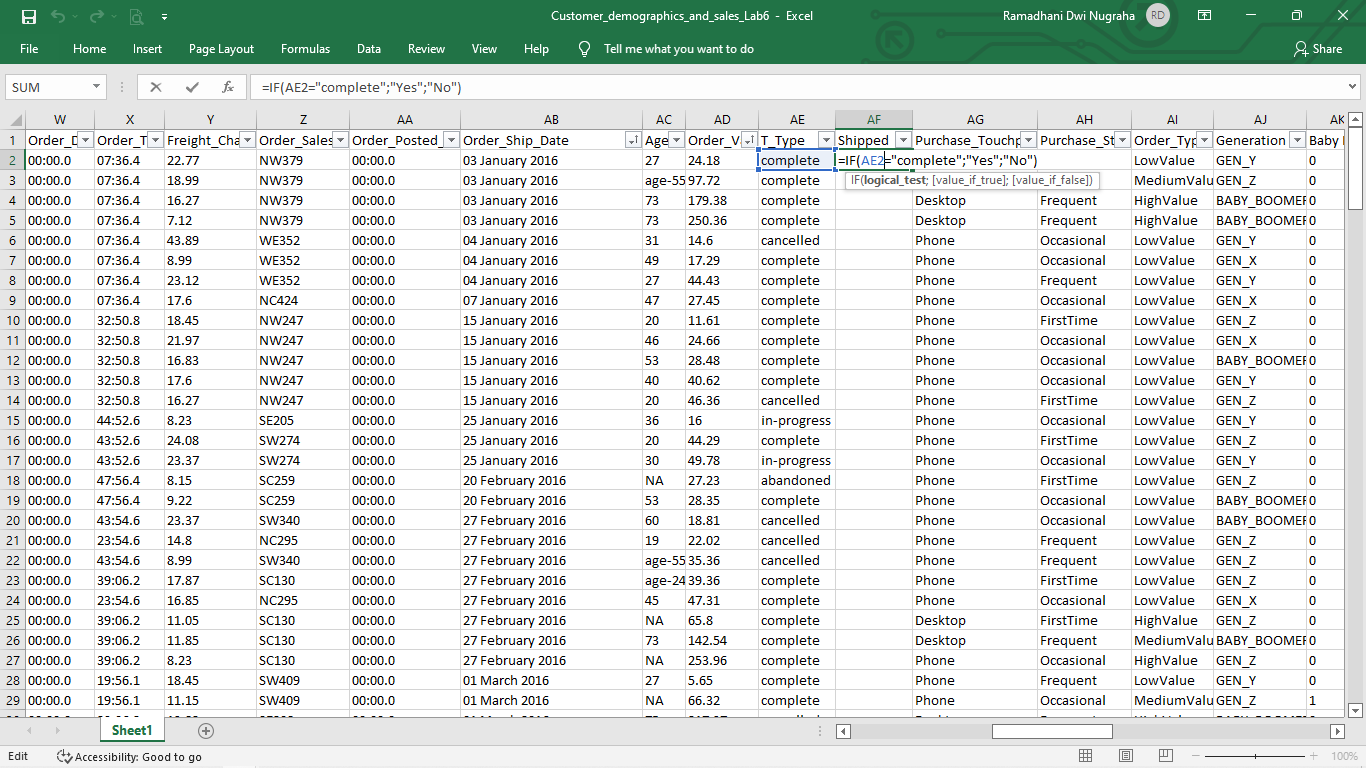
Setelah data sudah siap, maka kita dapat melakukan analisis untuk menemukan informasi lebih lanjut dari data tersebut. Dari bermacam-macam analisis yang bisa dilakukan, project ini melakukan analisis sederhana dari data tersebut dengan Filtering, Sorting, dan Fungsi.

Kita dapat melakukan sorting data dengan level yang berbeda. Hal ini dapat dilakukan dengan cara memilih custom sort pada Tab Data kemudian pilih Sort&Filter. Pada project ini dilakukan sorting data dengan mengurutkan Order\_Ship\_Date dari yang lama hingga terbaru dan Order\_Value dari yang terkecil hingga ke yang besar.

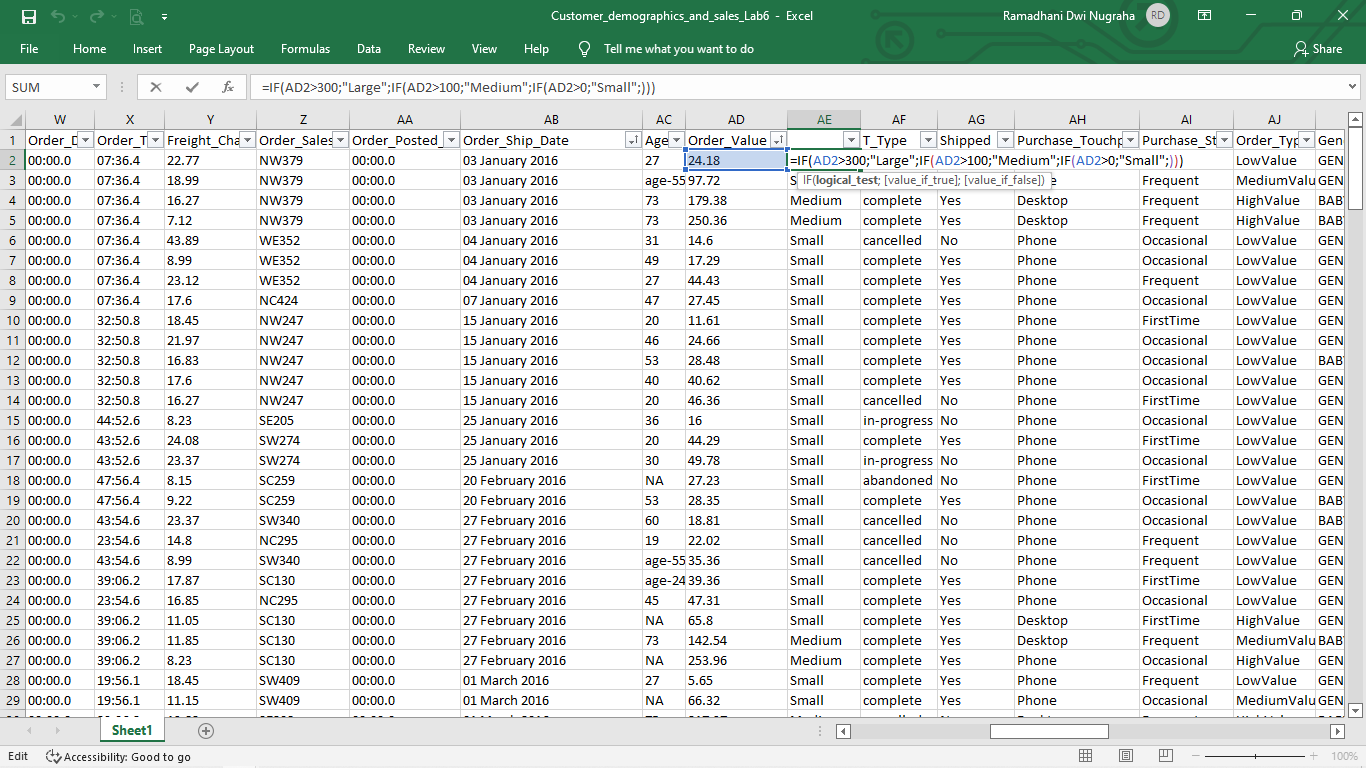
Data akan otomatis akan tersusun sesuai parameter yang kita berikan. Selanjutnya kita dapat mengaplikasikan beberapa fungsi untuk membantu kita dalam menganalisis data. Pertama yaitu fungsi IF. Fungsi ini memungkinkan Anda membuat perbandingan logis antara nilai dan perkiraan. Oleh karena itu, pernyataan IF dapat memiliki dua hasil. Hasil pertama jika perbandingan Anda Benar dan hasil kedua jika perbandingan Salah. Rumus fungsi ini memiliki sintaks:

=IF(logical\_test;[value\_if\_true]; [value\_if\_false])

Pada project ini fungsi IF digunakan untuk menyatakan barang sudah diantarkan atau belum berdasarkan kolom AE yaitu T\_Type.



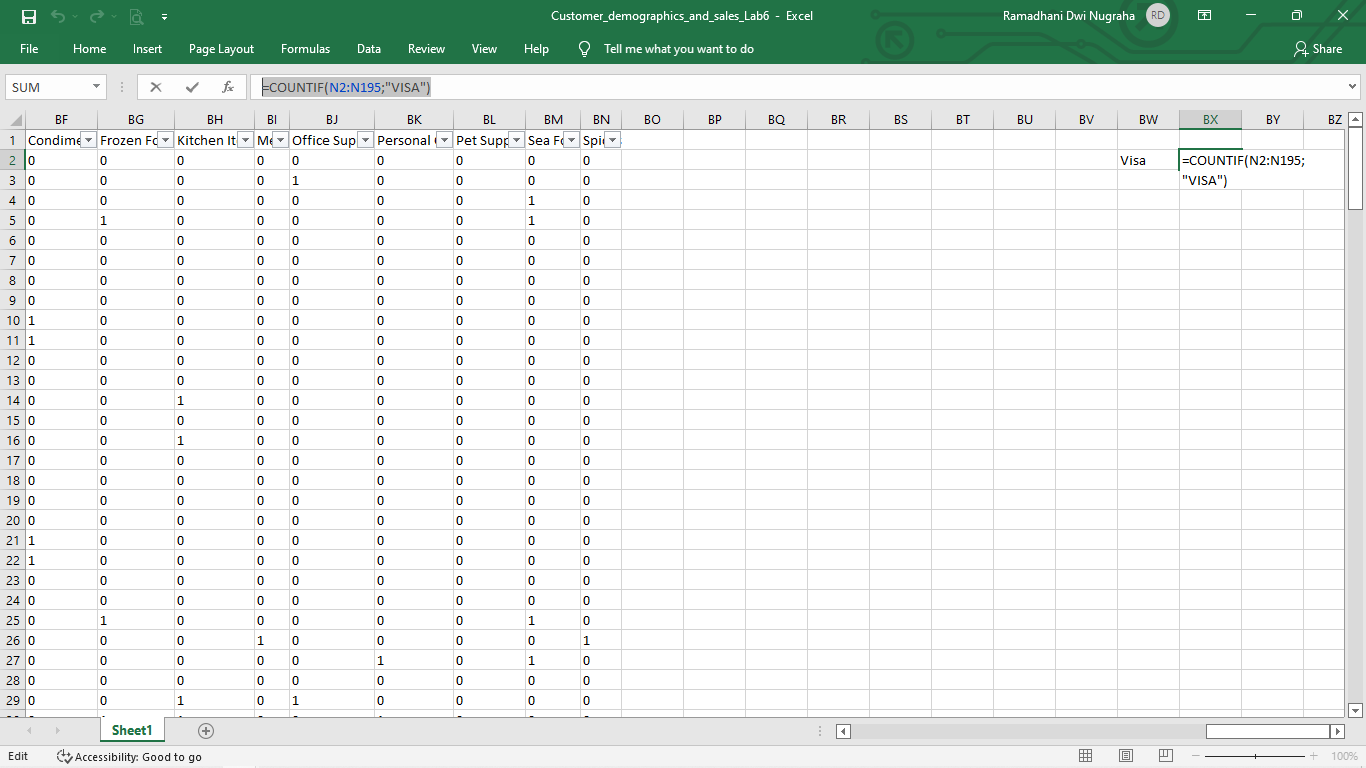
Fungsi IF juga dapat dilakukan untuk penggunaan beberapa kriteria seperti contohnya project ini akan mengkategorikan ukuran order menjadi Small, Medium, Large. Dengan batasan 0-100,100-300, Lebih dari 300.



Selain fungsi IF, terdapat pula fungsi yang cukup powerful dalam memberikan informasi terhadap data yang dimiliki yaitu COUNTIF. Fungsi COUNTIF digunakan untuk menghitung jumlah cell di worksheet dengen kriteria tertentu. Adapun sintaks dari Fungsi COUNTIF sendiri adalah

=COUNTIF(range; criteria)

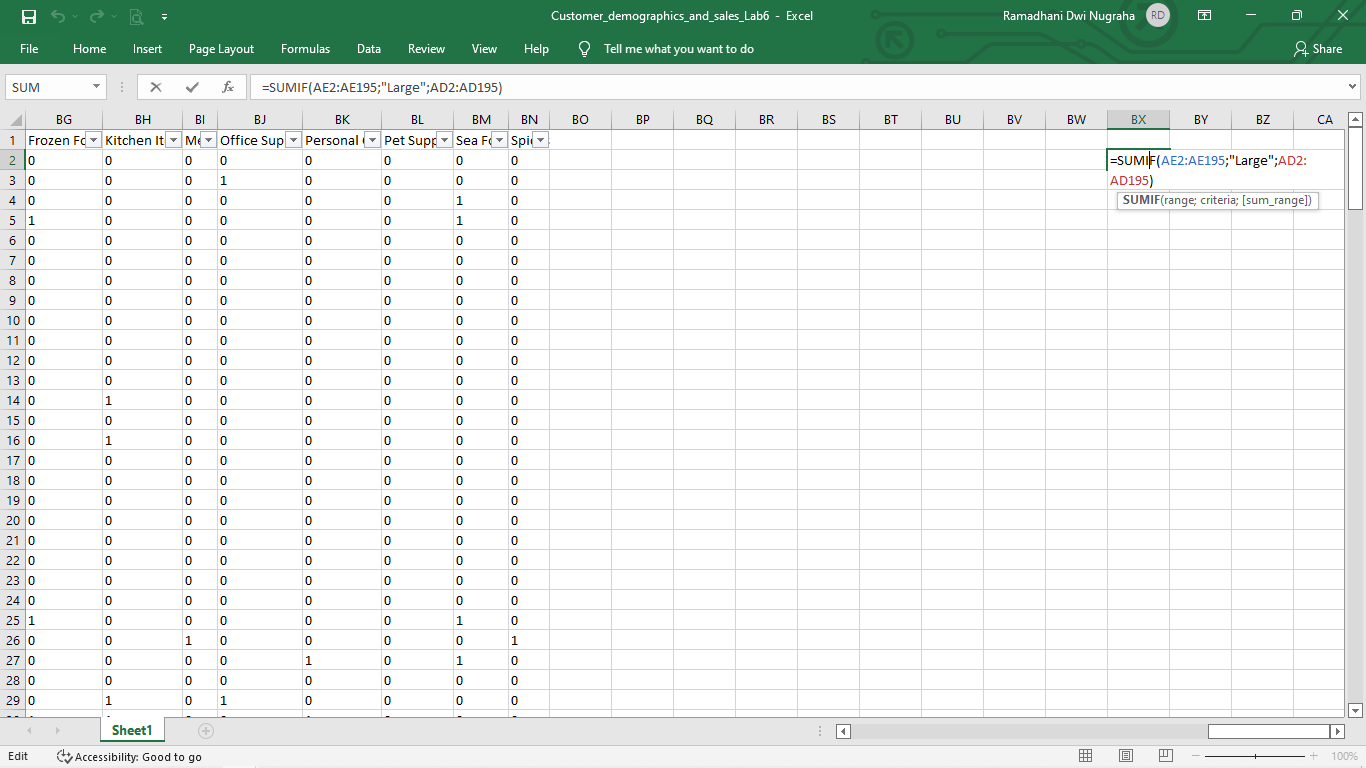
Pada project ini digunakan COUNTIF untuk menghitung jumlah cell dengan data berisikan “VISA” pada kolom N yaitu Creditcard\_Type



Adapun jumlah dari data yang berisikan VISA adalah 41 cell. Selanjutnya fungsi yang cukup powerful dalam menganalisis adalah SUMIF. Fungsi ini digunakan untuk menjumlahkan isi cell dengan kriteria tertentu. Adapun sintaks dari Fungsi ini adalah

=SUMIF(range;criteria;[sum\_range]

Pada project ini digunakan SUMIF untuk menjumlahkan kolom AD yaitu Order\_Value dengan kriteria “Large”.



Jumlah dari order value pada parameter Large adalah 2638,9.

Fungsi selanjutnya adalah fungsi VLOOKUP dan HLOOKUP. Fungsi VLOOKUP dan HLOOKUP dapat diaplikasikan untuk memindahkan suatu data ke tabel lainnya dengan kata kunci tertentu. Pada project ini fungsi VLOOKUP digunakan untuk mencari data dari Adress, Phone Number, dan Email yang dibuat pada tabel sendiri

1. https://dataplatform.cloud.ibm.com/exchange/public/entry/view/f8ccaf607372882403a37d9019b3abf4?utm\_medium=Exinfluencer&utm\_source=Exinfluencer&utm\_content=000026UJ&utm\_term=10006555&utm\_id=NA-SkillsNetwork-Channel-SkillsNetworkCoursesIBMDeveloperSkillsNetworkDA0130ENSkillsNetwork20531059-2021-01-01 [↑](#footnote-ref-1)