**Extract, Transform and Load Group Project**

Group Members: *Omar Aziz, Wilton Kwong, Ayobami Folaranmi, and Hocine Makhlouf*

**What we did?**

1. Sources of data that was **extracted**:

A. Jodidata.org - Downloaded a csv file containing monthly oil production by country over the last 5 years, between 02/2014 to 02/2019. [http://www.jodidb.org/TableViewer/tableView.aspx?ReportId=9390](http://www.jodidb.org/TableViewer/tableView.aspx?ReportId=93904)

B. Alpha Vantage - A website that provides free APIs for realtime and historical data on stocks, commodity prices, forex, and currencies (02/2014 to 02/2019).

https://www.alphavantage.co/documentation/#

Jodiddata.org provided monthly oil production, while Alpha vantage provided historical brent crude oil prices, which is the most popular benchmark used for marketing and selling crude oil around the world.

2. **Transformation** steps/methodology:

The extracted data (oil production and brent crude oil prices) was brought into the pandas library in Python to eliminate redundant or null values and then displayed in a data frame for easier data manipulation and analysis. Data was then merged together for easy curation and more robust analysis.

3. **Loading** the data (relational or non-relational):

Finally, data was stored in a mySQL database; a relational database which will be made available for public access and future use. The table in the database will have columns showing and rows showing countries, crude oil prices, crude oil production and monthly dates.

**Why we did it?**

Users will be able to create insights on how oil price volatility impacts production and revenue by country.

Users can also see patterns or trends in oil supply and demand on a global scale.