**Title**: A Statistical Analysis of Undocumented Immigrants Apprehended & Deported in the U.S.

**Data Source**:

*Note: The Kaggle dataset used in the notebook originated from the Department of Homeland and Security. The links from both sources are displayed.*

* Kaggle:
  + Project: [Immigration Nation Analytics](https://www.kaggle.com/ekayfabio/immigration-nation-analytics)
    - Dataset #1: [Undocumented Immigrants Apprehended in the U.S.](https://www.kaggle.com/ekayfabio/immigration-apprehended)
    - Dataset #2: [Undocumented Immigrants Deported in the U.S.](https://www.kaggle.com/ekayfabio/immigration-deported)
* Department of Homeland Security:
  + Publication Library: [Yearbook of Immigration Statistics 2019](https://www.dhs.gov/immigration-statistics/yearbook/2019?topics=all)
    - Dataset #1: [Aliens Apprehended: Fiscal Years 1925 to 2019](https://www.dhs.gov/immigration-statistics/yearbook/2019/table33)
    - Dataset #2: [Aliens Removed or Returned: Fiscal Years 1892 to 2019](https://www.dhs.gov/immigration-statistics/yearbook/2019/table39)

**Data Description**:

* What is your dataset about?
  + The apprehended dataset is number of undocumented immigrants arrested/seized in the U.S. for each year.
  + The deported dataset is the number of undocumented immigrants deported from the U.S. for each year. This dataset has two columns: (1) removals and (2) returns.
    - Removals: The deportation of undocumented immigrants based on an order of removal. Thus, the deportation is by force.
    - Returns: The deportation of undocumented immigrants not based on an order of removal. Thus, the deportation is likely by choice.
* What is the length of your data? Make sure you have enough data to build a model. *~100 Years*
  + The apprehended dataset ranges 94 years (94 points).
  + The deported (removed/returned) dataset ranges 127 years (127 points).
* What is the frequency of your data? Monthly, yearly? *Yearly*