Capstone Project Ideas

Number of jobs held

I found a great dataset from the US Bureau of labor statistics where they survey 9,964 men and women ages 14 to 22 when first interviewed in 1979 and ages 49 to 58 when interviewed most recently in 2014-15. The survey spans 35 years and provides information on work and nonwork experiences, education, training, income and assets, health, and other characteristics. The information provided by respondents, who were interviewed annually from 1979 to 1994 and biennially since 1994, can be considered representative of all men and women born in the late 1950s and early 1960s and living in the United States when the survey began in 1979.

Sample questions:
1) How many jobs does the average American hold from ages 18 - 58?
2) Does a persons education affect the number of jobs?
3) Does it affect the length of employment or unemployment?
4) What kind of role does gender or race play in these findings?

Link to data: https://www.bls.gov/nls/y79supp.htm

Top Spotify Tracks of 2017

At the end of each year, Spotify compiles a playlist of the songs streamed most often over the course of that year. This year's playlist (Top Tracks of 2017) included 100 songs. The audio features for each song were extracted using the Spotify Web API and the spotipy Python library. The dataset features audio features for each song, including danceability, tempo, key, etc.

Sample questions:
1) What were the top artists, genres, and audio features of the top 100 songs?
2) Do the top 100 or top 10 songs have anything in common?
3) Based on the audio features available in the data, is there a way to predict possible characteristics for future top 100 tracks?

Link to Data: https://www.kaggle.com/nadintamer/top-tracks-of-2017

Online Retail Purchases

This dataset contains all the transactions occurring between 01/12/2010 and 09/12/2011 for a UK-based online retailer. The company mainly sells unique all-occasion gifts.

Sample questions:
1) Are there any patterns in customers purchase choices?
2) Are there many 1st time buyers or return customers?
3) How do annual, monthly, weekly, and daily sales compare?
4) Are there specific times that have higher purchases? (i.e. Time of day, day of the week, Closer to the holidays, etc.)

Link to Data: http://archive.ics.uci.edu/ml/datasets/online+retail

Capstone Project Ideas

E-commerce Click Events & Purchases
Given a sequence of click events performed by a user during a typical session in an e-commerce website, the goal is to predict whether the user is going to buy something or not, and if he is buying, what would be the items he is going to buy.

Sample questions:
1) Is it possible to predict if a user will purchase something based on their website clicks?
2) What are the click events that may determine if the clicks will turn into a purchase?
3) Is it possible to predict which items the user may purchase?
4) Based on these findings is it possible to

Link to Data: http://recsys.yoochoose.net/challenge.html