Project Idea 1:

Problem statement: Is it possible to predict sale price of residential properties based on housing characteristics?

Dataset: District of Columbia shares hundreds of databases with general public and various government agencies on opendata.dc.gov website. One such dataset is Residential CAMA Computer Assisted Mass Appraisal database.
The dataset has 107K records and 39 variables which are various attributes of housing characteristics along with sale price.

Link : http://opendata.dc.gov/datasets/computer-assisted-mass-appraisal-residential

Techniques: Initial analysis indicates below techniques/methods may be needed to come up with predictive model.

- Exploratory Data analysis.
- Visualization
- Data imputation
- Regression
- Outlier analysis

Project Idea 2:

Problem statement: Is it possible to forecast the wholesale price index in India? Is there a relationship between consumer price index and wholesale price index?

Dataset: Government of India publishes monthly WPI. The current series has a base year of 2011-12. The dataset has around 74 monthly data across 800+ underlying items. Also it provides consumer price index.

Link: https://data.gov.in/resources/wholesale-price-index-base-year-2011-12-upto-may-2017

Techniques: Initial analysis indicates below techniques/methods may be needed to come up with predictive model.

- Data wrangling. Data is in report format. It needs to be formatted to tabular format ready for data science analysis.
- Exploratory Data analysis.
- Visualization
- Data imputation
- Time series forecasting and correlation.

Project Idea 3:

Problem statement: Is there a pattern of housing and neighborhood condition based on personal attributes such as race, disability or military personnel?

Dataset: In the AHS microdata, the basic unit is an individual housing unit. Each record shows most of the information associated with a specific housing unit or individual, except for data items that could be used to personally identify that housing unit or individual.

Link:
https://www.census.gov/programs-surveys/ahs/data/2015/ahs-2015-public-use-file--puf-/ahs-2015-national-public-use-file--puf-.html

Techniques: Initial analysis indicates below techniques/methods may be needed to come up with predictive model.

- Data wrangling. It is huge data. ~170 MB.
- Exploratory Data analysis.
- Visualization
- Data imputation
- Clustering