

## **IRisk Lab Data Discovery and Consolidation**

### **Week 11 Report**

*Rohit Valmeekam*

#### **What Was Accomplished**

- ❖ The primary emphasis was on the evaluation of identified datasets, further documentation, and an exploration of additional external resources to enhance our understanding of insurance premiums, pre-existing conditions, and health insurance market trends.
- ❖ Dataset Evaluation Progress:
  - Continued the evaluation process of the identified datasets,
    - This involved a thorough examination of metadata, scrutiny of data quality, and gaining insights into the variables present in each dataset.
- ❖ Integration with Data Pipeline:
  - Furthered the integration of data categories from the identified datasets into the data pipeline established earlier. This step is crucial for facilitating a smooth flow of information and laying the groundwork for future stages of the project.
- ❖ Documentation Refinement and Enrichment:
  - Continued documenting key features and attributes of each dataset
    - Including details on data sources, update frequency, and any notable limitations on spreadsheet
    - Emphasized the ease of incorporating the identified datasets related to insurance premiums, pre-existing conditions, and health insurance market trends into the existing data pipeline and extending them into the evolving database schema.

#### **Next Steps**

- ❖ The ongoing dataset evaluation process will be completed, addressing any outstanding considerations and ensuring that only high-quality, relevant datasets are incorporated into the project.