



Weekly Capstone Update Meeting

Nov 10 2022

PA Dept. of Labor & Industry + CMU

Goals for today

- Discuss work completed this past week
- Discuss hosting on IIS and adding new features to web application
- Discuss the visualization of analysis reports
- Discuss the progress on the severity calculation
- Discuss action items for next weeks

Done this past week

November/December

23 Sprint 4

30 Sprint 5

6 Sprint 6

13 Sprint 7

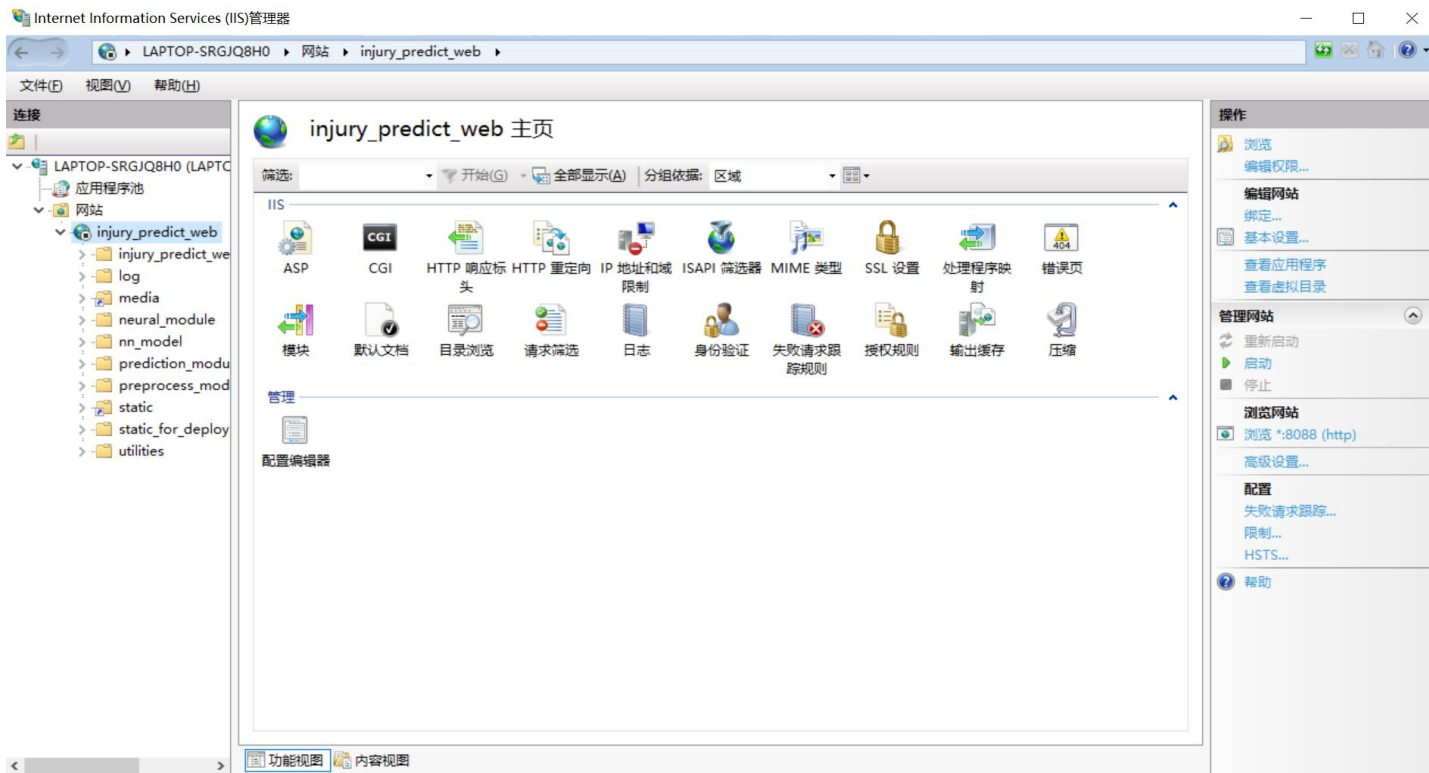


20 Finalize Build & Break

- Successfully run the the application on IIS and updated onboarding tutorials
- Further data preparation for prediction models: classify & regress medical costs
- Research about methods to define and calculate severity
- Format and optimize the UI layout of the Power BI report

Hosting the web application on IIS

Run the application on IIS





Missing medical costs

Missing medical costs

- Of claims with medical costs, $< 1\%$ have \$0
 - Less worried about classifying claims into medical cost vs. no medical cost
- Tested imputation (filling missing costs) methods
 - Linear Regression
 - Decision Trees
 - Random Forest
 - XGBoost
 - Still tuning and will choose best method next week.

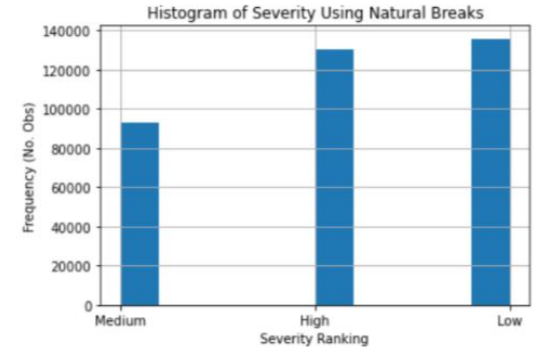
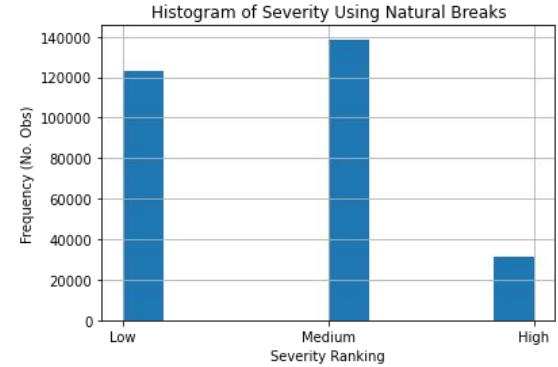
Severity Calculation

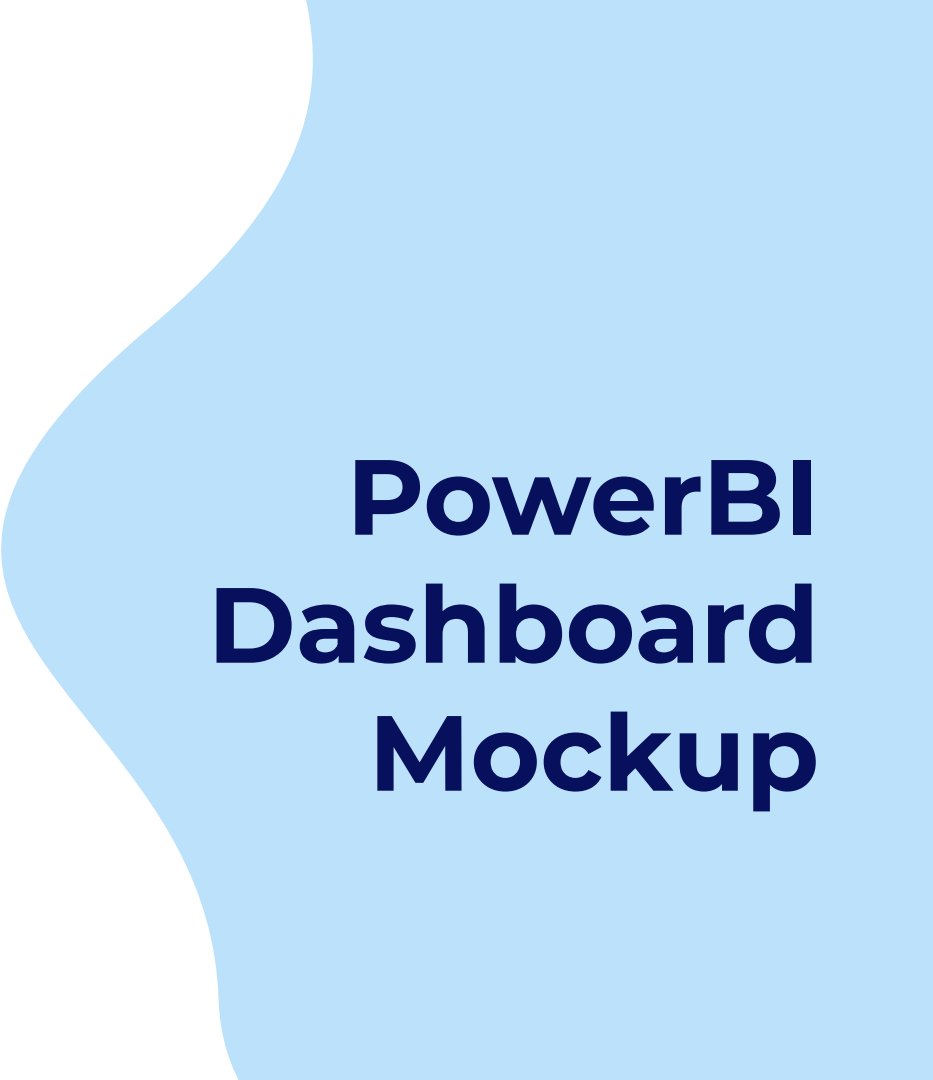
Defining severity

Explored different methods to assign severity with jenks natural breaks

Going to use a method that groups medical costs with injury and nature codes and use natural breaks to bucket severity

Examples of different breaks with same algorithm →





PowerBI Dashboard Mockup

Mockup dashboard

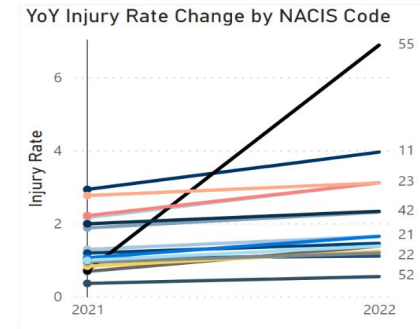
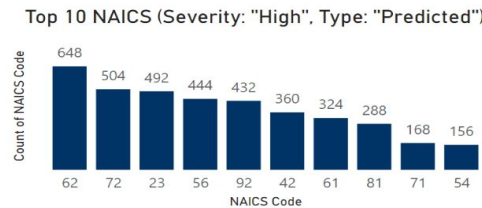
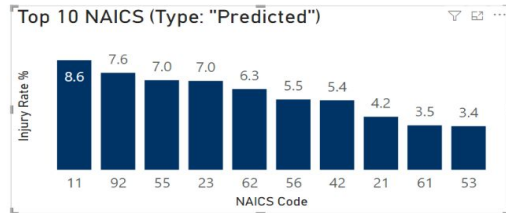


Injury Rate Analysis

Task ID: 23 | NAICS Level: All | NAICS Code: All | County: All | Severity: All

Previous Year Injury Rate
0.199 %
93,188 Injured / 46,870,766 Jobs

Next Year Injury Rate
0.238 %
113,914 Injured / 47,765,088 Jobs



Heat Map of NAICS by County Name

| naics_code | Adams County | Allegheny County | Armstrong County | Beaver County | Bedford County | Berks County | Blair County | Bradford County | Bucks County | Butler County | Cambria Co |
|------------|--------------|------------------|------------------|---------------|----------------|--------------|--------------|-----------------|--------------|---------------|------------|
| 11 | 0.10 | 0.06 | 0.08 | 0.13 | 0.05 | 0.12 | 0.05 | 0.04 | 0.06 | 0.09 | |
| 21 | | 0.01 | 0.12 | 0.04 | 0.00 | 0.12 | | 0.05 | 0.07 | 0.11 | |
| 22 | | 0.05 | 0.08 | 0.01 | 0.03 | 0.04 | 0.00 | | 0.03 | 0.04 | |
| 23 | 0.10 | 0.06 | 0.06 | 0.04 | 0.06 | 0.09 | 0.06 | 0.05 | 0.06 | 0.07 | |
| 42 | 0.04 | 0.07 | 0.12 | 0.04 | 0.02 | 0.07 | 0.06 | 0.02 | 0.04 | 0.06 | |
| 51 | | 0.03 | 0.01 | 0.28 | 0.02 | 0.04 | 0.02 | 0.01 | 0.02 | 0.03 | |
| 52 | 0.01 | 0.00 | 0.01 | | 0.06 | 0.02 | 0.01 | 0.00 | 0.01 | 0.01 | |
| 53 | 0.01 | 0.05 | 0.05 | 0.03 | | 0.06 | 0.04 | 0.07 | 0.05 | 0.03 | |
| < | | | | | | | | | | | > |

Still in progress, numbers are incorrect.

Next Week's Sprint

- Implement new severity definition
- Implement missing medical costs imputations
- Format and optimize the UI layout of the Power BI report
- Build onboarding documentation