# Week 5

Advay Vyas

5/6/25

## Contents



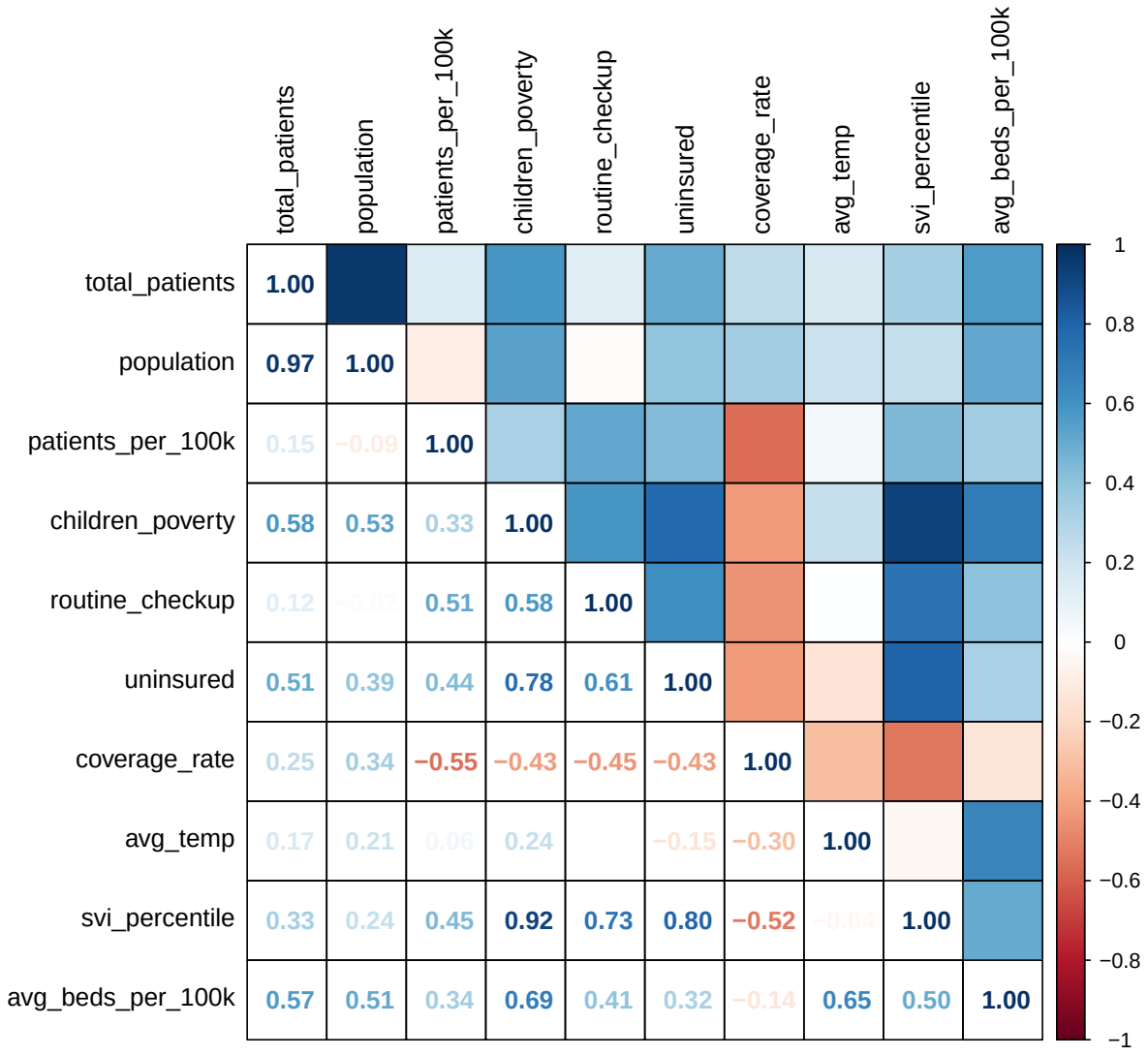
---

## Introduction

This week, we are going to conduct multiple linear regression and try to incorporate interaction variables and ANOVA for some additional modeling boosts. The (tentative) indicator variables are flu vaccination coverage rate and uninsured % (from Week 4), temperature (from Week 3), Social Vulnerability Index (SVI) and beds per capita (from Week 2). Since this is my last week until the break ends, I wanted to conduct a sort of final project that would try to tie everything together and hopefully show some real trends from the data that I've web scraped and cleaned.

## The old model

## A new correlation matrix




## The new model

### Adding flu vaccination coverage rate

| term | estimate | std_error | statistic | p_value | lower_ci | upper_ci |
|---|---|---|---|---|---|---|
| intercept | 141.407 | 53.000 | 2.668 | 0.037 | 11.720 | 271.093 |
| coverage_rate | -2.196 | 1.352 | -1.625 | 0.155 | -5.504 | 1.112 |

```
## [1] 0.3055056
```

| term | estimate | std_error | statistic | p_value | lower_ci | upper_ci |
|---|---|---|---|---|---|---|
| intercept | 141.407 | 53.000 | 2.668 | 0.037 | 11.720 | 271.093 |
| coverage_rate | -2.196 | 1.352 | -1.625 | 0.155 | -5.504 | 1.112 |

```
## [1] 0.5595952
```

### Adding routine checkup coverage percentage

| term | estimate | std_error | statistic | p_value | lower_ci | upper_ci |
|---|---|---|---|---|---|---|
| intercept | -39.473 | 217.555 | -0.181 | 0.863 | -598.718 | 519.771 |
| coverage_rate | -1.601 | 1.546 | -1.035 | 0.348 | -5.576 | 2.374 |
| routine_checkup | 2.189 | 2.550 | 0.858 | 0.430 | -4.365 | 8.743 |

```
## [1] 0.3947241
```

### Adding temperature data

| term | estimate | std_error | statistic | p_value | lower_ci | upper_ci |
|---|---|---|---|---|---|---|
| intercept | -9.998 | 288.909 | -0.035 | 0.974 | -812.138 | 792.143 |
| coverage_rate | -1.716 | 1.828 | -0.939 | 0.401 | -6.791 | 3.359 |
| routine_checkup | 2.106 | 2.872 | 0.733 | 0.504 | -5.867 | 10.080 |
| avg_temp | -0.276 | 1.476 | -0.187 | 0.861 | -4.374 | 3.821 |

```
## [1] 0.3999752
```

### Adding SVI data

| term | estimate | std_error | statistic | p_value | lower_ci | upper_ci |
|---|---|---|---|---|---|---|
| intercept | -17.018 | 355.575 | -0.048 | 0.965 | -1148.617 | 1114.580 |
| coverage_rate | -1.763 | 2.263 | -0.779 | 0.493 | -8.965 | 5.439 |
| routine_checkup | 2.264 | 4.320 | 0.524 | 0.637 | -11.485 | 16.013 |
| avg_temp | -0.295 | 1.734 | -0.170 | 0.876 | -5.815 | 5.225 |
| svi_percentile | -1.779 | 31.308 | -0.057 | 0.958 | -101.416 | 97.858 |

```
## [1] 0.4006205
```

### Adding beds per capita

| term | estimate | std_error | statistic | p_value | lower_ci | upper_ci |
|---|---|---|---|---|---|---|
| intercept | 653.503 | 370.954 | 1.762 | 0.220 | -942.582 | 2249.587 |
| coverage_rate | -6.241 | 2.434 | -2.564 | 0.124 | -16.713 | 4.232 |
| routine_checkup | 0.896 | 2.839 | 0.316 | 0.782 | -11.321 | 13.114 |

| term | estimate | std_error | statistic | p_value | lower_ci | upper_ci |
|---|---|---|---|---|---|---|
| avg_temp | -6.470 | 2.913 | -2.221 | 0.156 | -19.003 | 6.063 |
| svi_percentile | -69.484 | 35.713 | -1.946 | 0.191 | -223.147 | 84.178 |
| avg_beds_per_100k | 0.240 | 0.105 | 2.295 | 0.149 | -0.210 | 0.691 |

```
## [1] 0.8349919
```

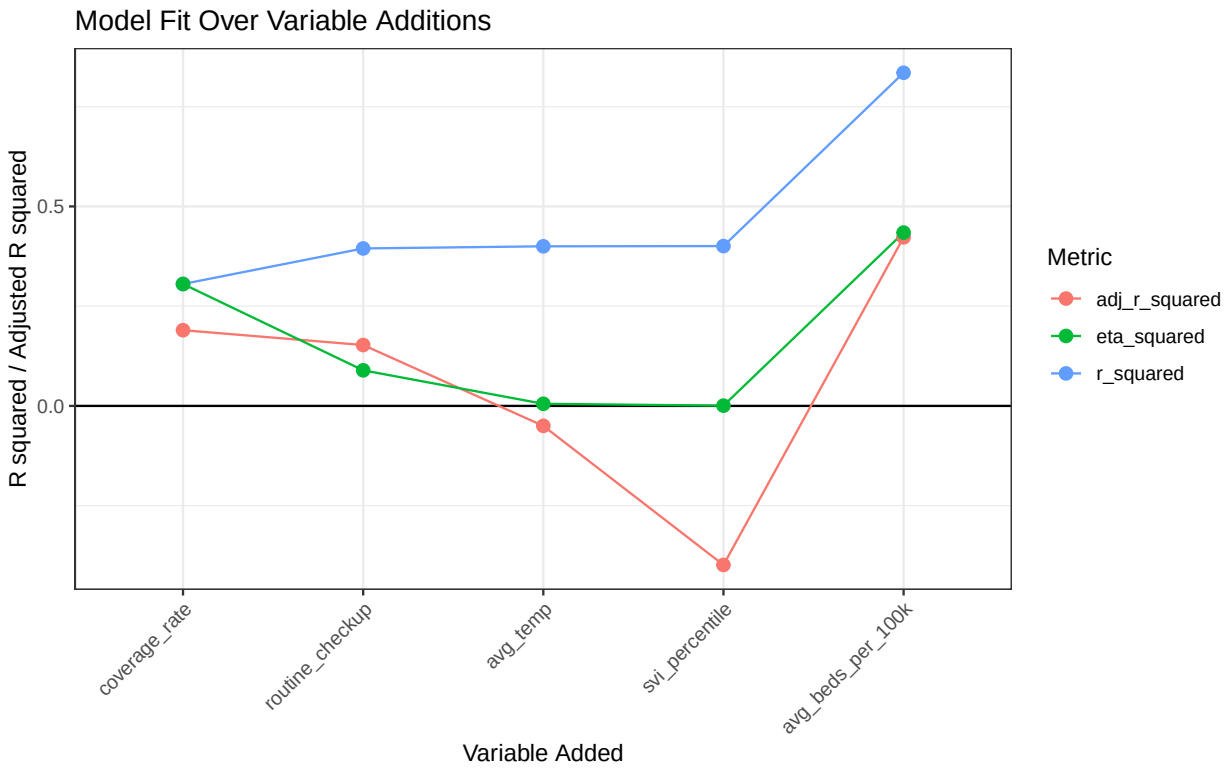
# ANOVA

Table 7: $R^2$ table

| Variable | $R^2$ | Adjusted $R^2$ |
|---|---|---|
| coverage_rate | 0.3055 | 0.1898 |
| routine_checkup | 0.3947 | 0.1526 |
| avg_temp | 0.4000 | -0.0500 |
| svi_percentile | 0.4006 | -0.3986 |
| avg_beds_per_100k | 0.8350 | 0.4225 |

Table 8: $\eta^2$ table

| Variable | $\eta^2$ | Confidence | Lower | Upper |
|---|---|---|---|---|
| coverage_rate | 0.3055056 | 0.95 | 0 | 1 |
| routine_checkup | 0.0892185 | 0.95 | 0 | 1 |
| avg_temp | 0.0052512 | 0.95 | 0 | 1 |
| svi_percentile | 0.0006452 | 0.95 | 0 | 1 |
| avg_beds_per_100k | 0.4343714 | 0.95 | 0 | 1 |




### Results

### Analysis

## Conclusion

### Week 5

### Looking forward

What a ride! For how short this was, it was honestly great.