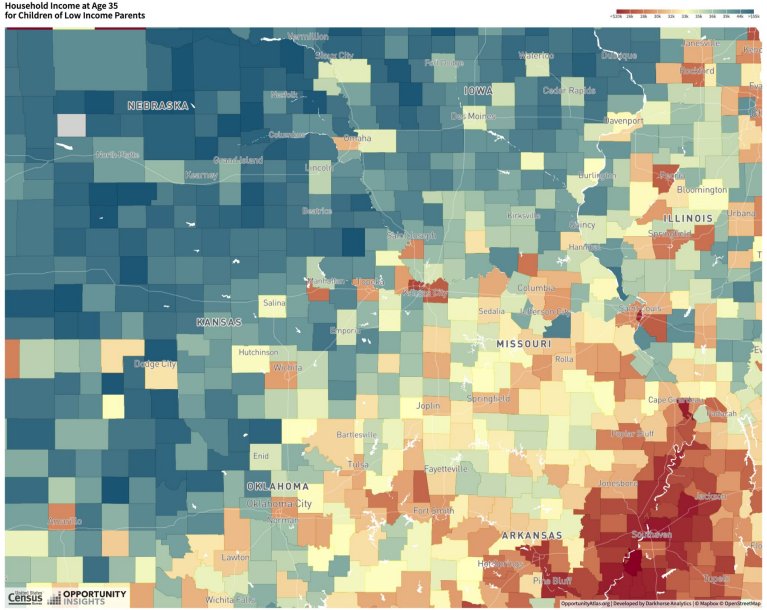
# The relationship between Absolute upward mobility at 25% and innovation rate

Keing Zhang

# KS vs MO

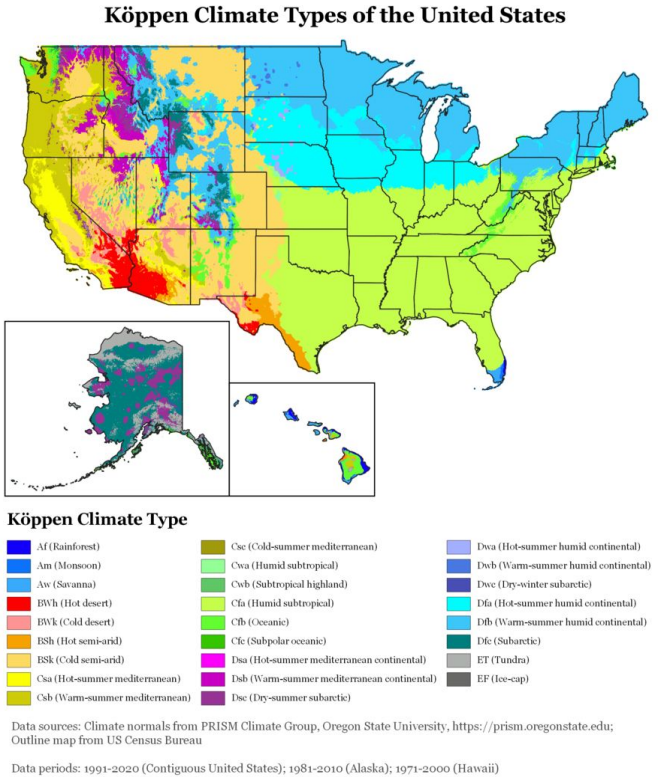



What causes the difference?

- climate
- Race and ethnicity
- Population

# Climate

- Similar Climate



# Race

## KS

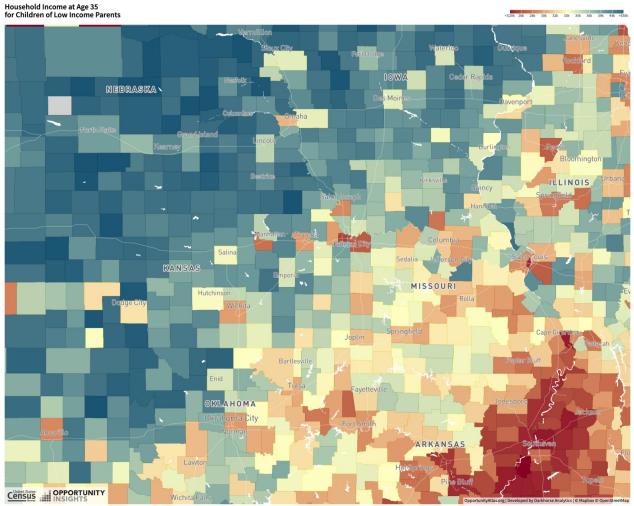
White alone 61.6%; Black alone 12.4%; Hispanic 18.7%; Asian alone 6%; American Indian and Alaska Native alone 1.1%; Native Hawaiian and Other Pacific Islander alone 0.2%; Some Other Race alone 8.4%; Two or More Races 10.2%

## MO

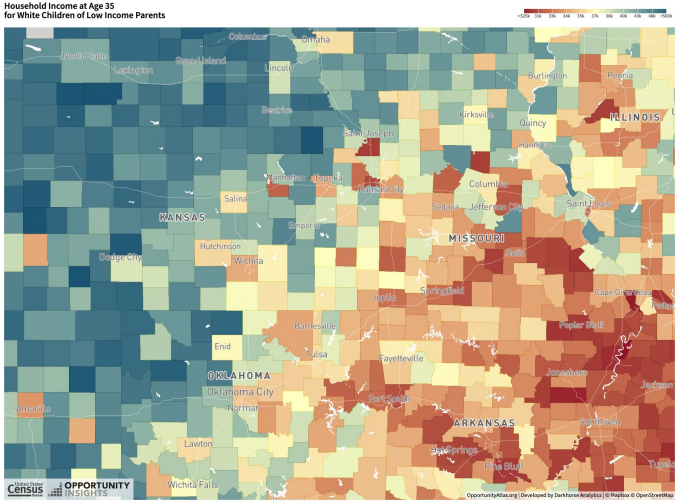
White alone 61.6%; Black alone 12.4%; Hispanic 18.7%; Asian alone 6%; American Indian and Alaska Native alone 1.1%; Native Hawaiian and Other Pacific Islander alone 0.2%; Some Other Race alone 8.4%; Two or More Races 10.2%

Source: https://www.census.gov/library/stories/state-by-state/kansas-population-change-between-census-decade.html#:~:text=Population%20(up%207.4%25%20to%20331.4,or%20More%20Races%2010.2%25).
https://www.census.gov/library/stories/state-by-state/missouri-population-change-between-census-decade.html#:~:text=Population%20(up%207.4%25%20to%20331.4,or%20More%20Races%2010.2%25).

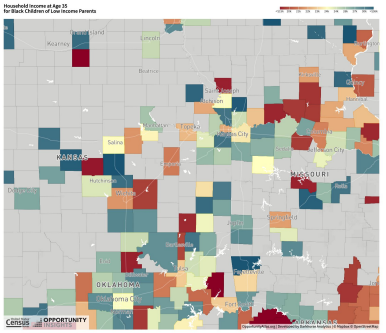
# Income at 35 divided by race



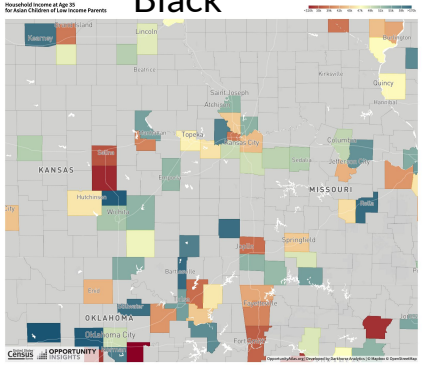

All race




White

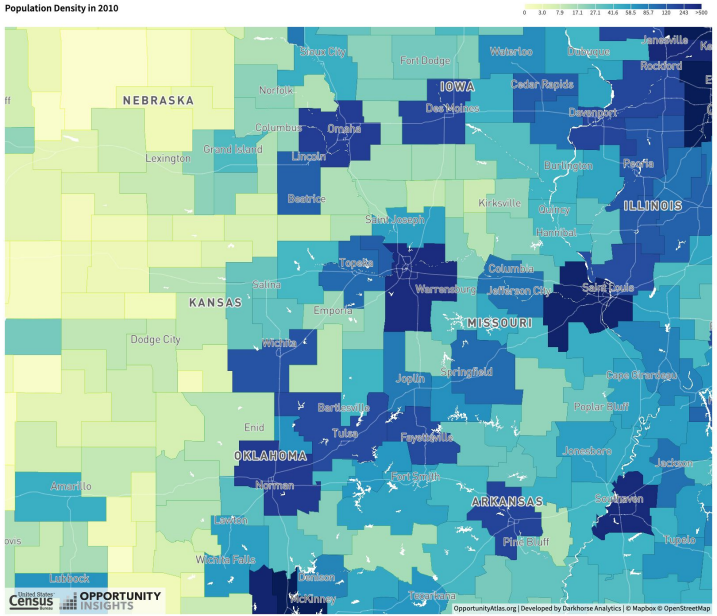



Black



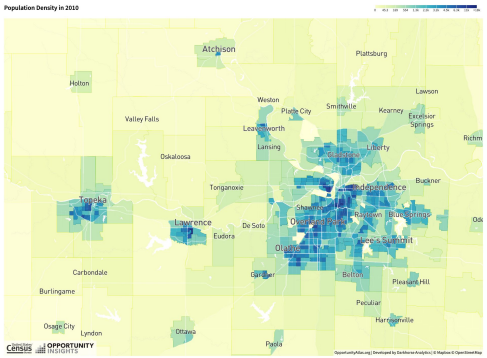

Asian

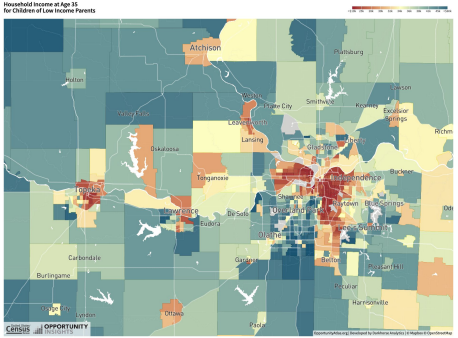
# Population density



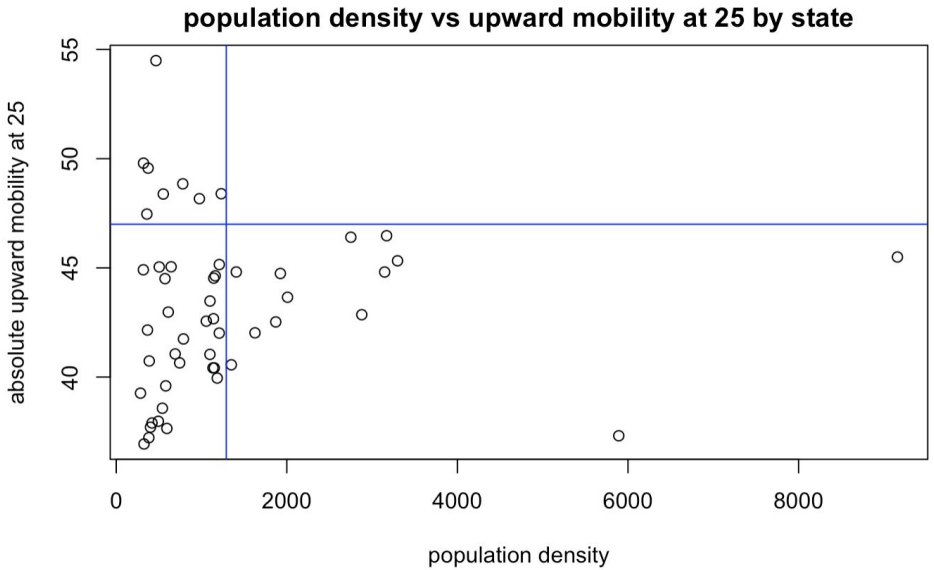

Kansas: 2.913 million/82,278 mi²

Missouri: 6.124 million/69,715 mi²



# Visualization




Population density is not related to upward mobility

r=0.0576475

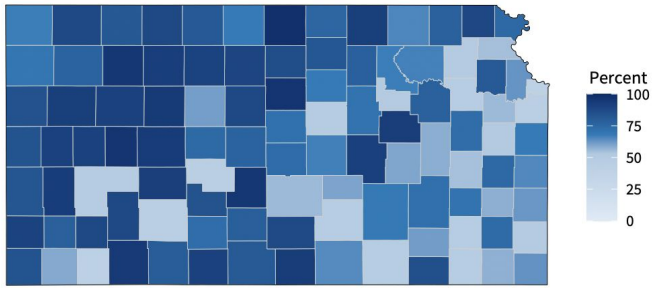
There's no state with high population density and high upward mobility.

# Small business

## Kansas

**258,012** small businesses
**99.1 percent** of Kansas businesses

**601,550** small business employees
**50.0 percent** of Kansas employees

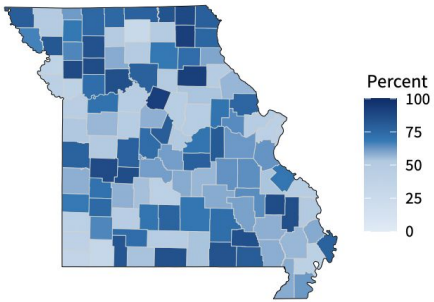



**Share of employees working at small businesses by county**
Source of data: Statistics of US Businesses (Census)

## Missouri

**542,519** small businesses
**99.4 percent** of Missouri businesses

**1.2 million** small business employees
**45.8 percent** of Missouri employees




**Share of employees working at small businesses by county or independent city**
Source of data: Statistics of US Businesses (Census)

https://advocacy.sba.gov/2021/08/30/2021-small-business-profiles-for-the-states-the-district-of-columbia-and-the-u-s/

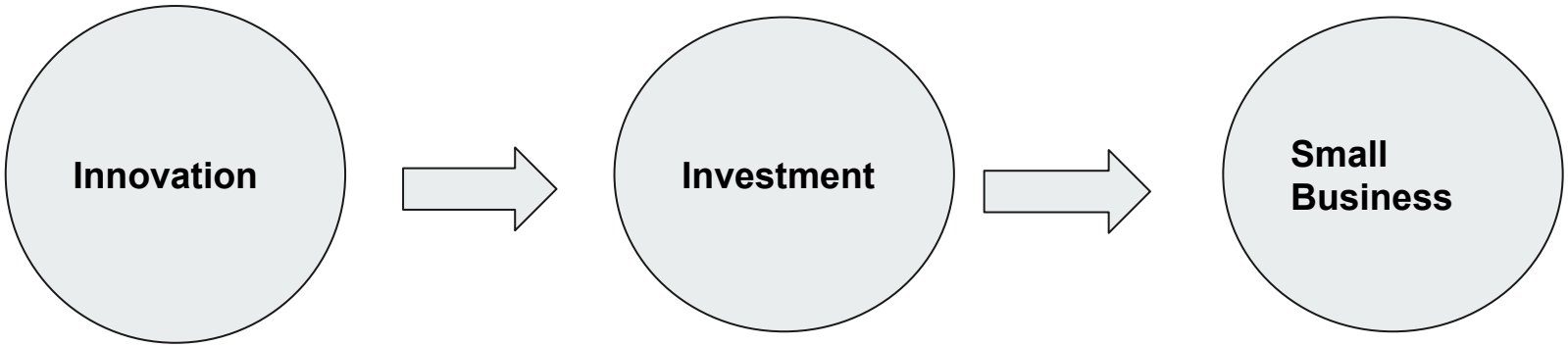
# What happens here?

As there are more than 2 times more people in Missouri, it is harder to get more resources for individuals in the same group(region/ race, etc). Under this background, compare to "**innovation**"(more cost), it is safer to do "**imitation**" as an individual, which leads to "**homogenous competition.**"

Thus, due to homogenous competition dominated by large company, it is harder to get high salary.

If there are more small business( usually started by innovation), it is easier to get high salary due to less "homogenous competition."

# How innovations gain profits



# Hypothesis:

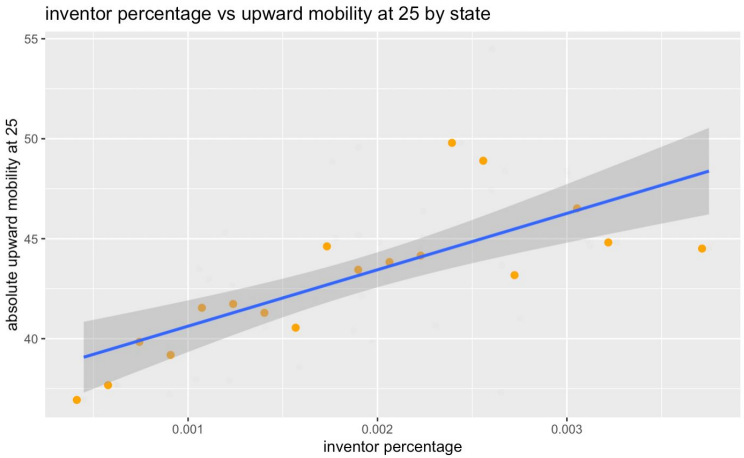
Upward mobility is positively related to innovation percentage  which leads to profitable small businesses.

(Outside Dataset: Inventors in America: Commuting Zone Innovation Rates by Childhood Commuting Zone, Gender, and Parent Income) merged with atlas by region

# Dataset

| Variable | Description |
|---|---|
| par_cz | Childhood commuting zone of residence |
| par_czname | Commuting zone name |
| par_state | Childhood state Federal Information Processing Standard (FIPS) code; CZs that cross state borders are assigned to the state which contains the largest population in the CZ, based on the 2010 Census |
| par_stateabbrv | Two-letter state abbreviation |
| kid_count | Number of children |
| inventor | Share of children who go on to become inventors |
| top5cit | Share of children with patent citations in top 5% of their birth cohort, using total number of citations |
| inventor_cat_[c] | Share of children who patent in technology category [c].<br>Technology categories [c] are:<br>1 – Chemical<br>2 – Computers and Communications<br>3 – Drugs and Medical<br>4 – Electrical and Electronic<br>5 – Mechanical<br>6 – Others<br>7 – Design and Plant |
| top5cit_cat_[c] | Share of children who patent in technology category [c] and have total patent citations in top 5% of their birth cohort |
| [outcome]_g_m | Identical to variable [outcome], but restricting the sample to males. |
| [outcome]_g_f | Identical to variable [outcome], but restricting the sample to females. |
| [outcome]_pq_[quintile] | Identical to variable [outcome], but restricting the sample to children whose parental income is in quintile [quintile] of the parent income distribution of the children's birth cohort. |

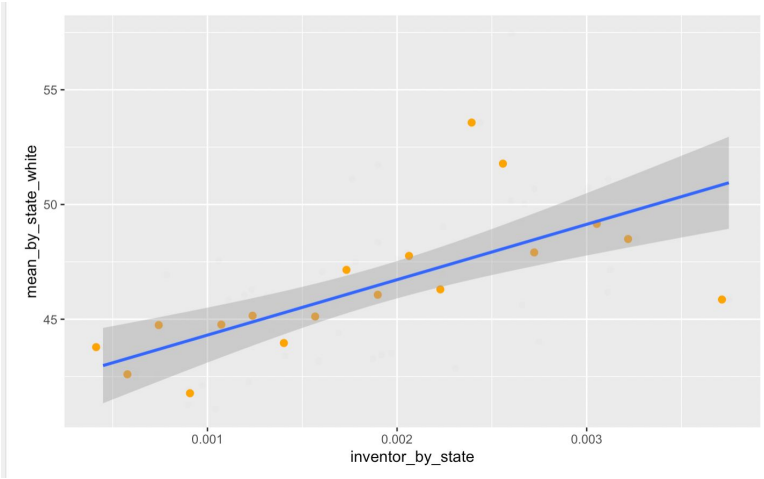
# Visualization

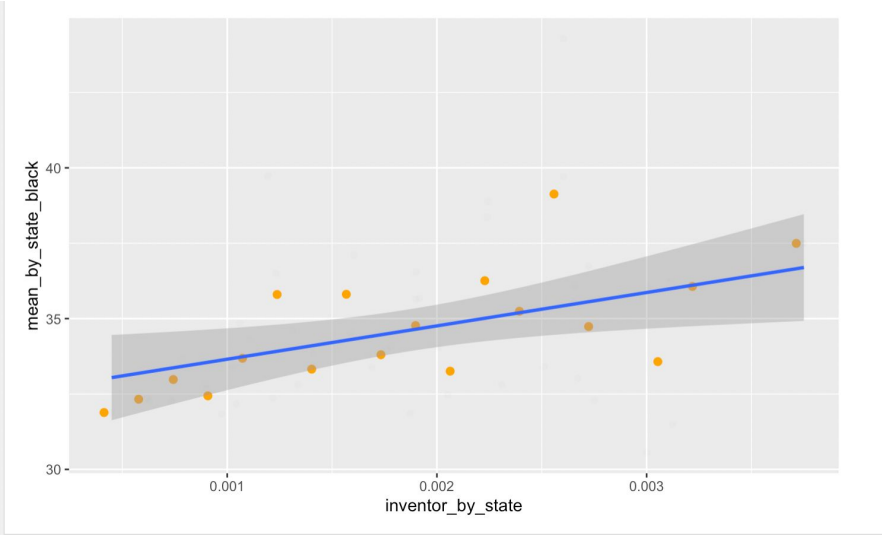



The inventor percentage is positively related to upward mobility at 25%.

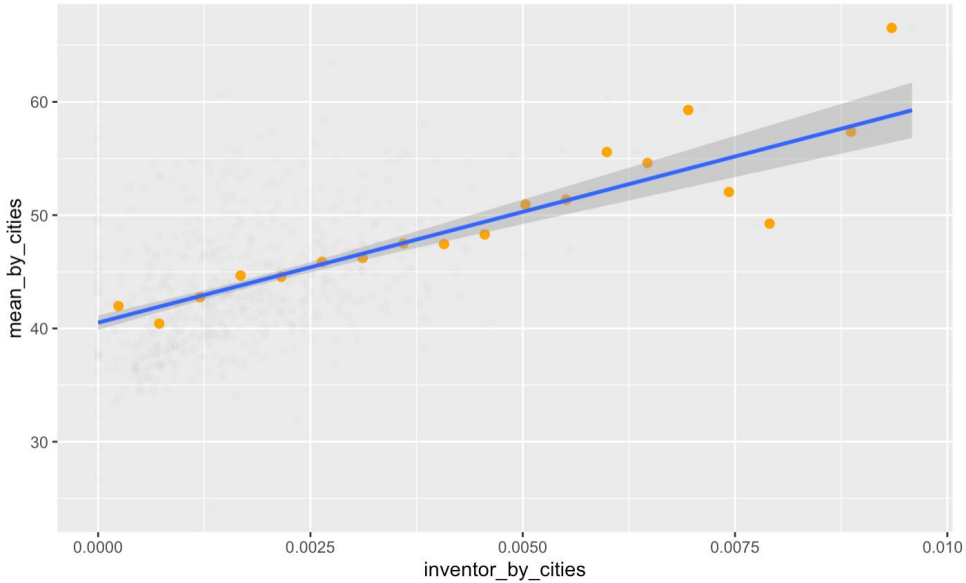
Correlation coefficient =0.6042194

# Visualization white/ black



# At city level



# Regression Model

```
Call:
lm(formula = res$mean_by_state ~ res$inventor_by_state + res$pop_density)

Residuals:
    Min      1Q  Median      3Q     Max
-6.7906 -2.1621 -0.0602  1.9458  8.9976

Coefficients:
                      Estimate Std. Error t value Pr(>|t|)
(Intercept)           3.791e+01  1.103e+00  34.389  < 2e-16 ***
res$inventor_by_state 2.962e+03  5.521e+02   5.365 2.31e-06 ***
res$pop_density      -2.835e-04  2.971e-04  -0.954    0.345
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 3.09 on 48 degrees of freedom
Multiple R-squared:  0.3769,    Adjusted R-squared:  0.3509
F-statistic: 14.52 on 2 and 48 DF,  p-value: 1.173e-05
```

Population density is not a statistically significant factor. So remove it.

# Regression Model

```
Call:
lm(formula = res$mean_by_state ~ res$inventor_by_state)

Residuals:
    Min      1Q  Median      3Q     Max
-7.9769 -1.9558 -0.1142  1.9367  9.3425

Coefficients:
                      Estimate Std. Error t value Pr(>|t|)
(Intercept)             37.806      1.096  34.506  < 2e-16 ***
res$inventor_by_state 2820.076    531.285   5.308 2.67e-06 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 3.088 on 49 degrees of freedom
Multiple R-squared:  0.3651,    Adjusted R-squared:  0.3521
F-statistic: 28.18 on 1 and 49 DF,  p-value: 2.669e-06
```

36.5% of the variability observed in the target variable is explained by the regression model.

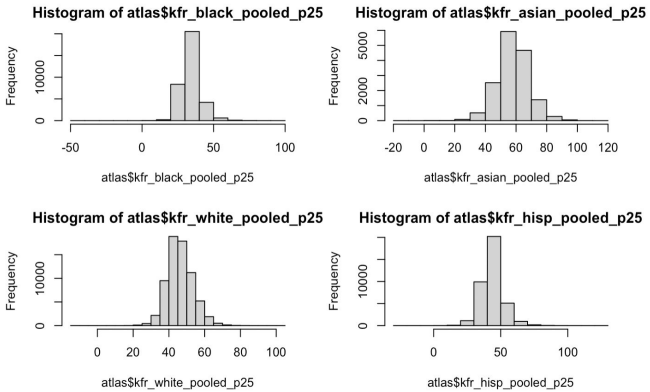
# Regression Model

```
Call:
lm(formula = res$mean_by_state ~ res$inventor_by_state + res$singleparent +
    res$poor)

Residuals:
    Min      1Q  Median      3Q     Max
-3.9471 -1.5718 -0.1011  1.3405  6.0872

Coefficients:
                     Estimate Std. Error t value Pr(>|t|)
(Intercept)            49.421      2.412  20.489  < 2e-16 ***
res$inventor_by_state 2644.423    465.325   5.683 8.10e-07 ***
res$singleparent       -51.275      7.923  -6.472 5.18e-08 ***
res$poor                29.215     13.850   2.109   0.0403 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 2.246 on 47 degrees of freedom
Multiple R-squared:  0.6777,    Adjusted R-squared:  0.6572
F-statistic: 32.95 on 3 and 47 DF,  p-value: 1.28e-11
```

67.77% of the variability observed in the target variable is explained by the regression model.

# Limitation

- Lack of data for Black, Asian, and Hispanic on the map
- Inventors data does not include race variable
- Failed to get small business dataset
- Inventions could be categorized into more categories
- It is too broad to compare at state level




Black: 33.98718

Asian: 57.9756

White: 46.29786

Hispanic: 43.70138

# Conclusion

There’s an association between inventor percentage and upward mobility at 25%. Yet, there’s not enough evidence to say that innovation percentage has a cause and effect relationship with  fraction of small businesses.

# Thank you!

Questions/ Comments/ Advice for future research