project_dv_pe7112

Project Title:

Analyzing the Impact of Rail and Air Transport on U.S. State Economies

Project Overview:

This project employs a data-driven approach to visualize and interpret the contributions of rail and air transport sectors to the Gross Domestic Product (GDP) of U.S. states.

Data Collection:

GDP data was collected for rail and air transportation sectors for each U.S. state. (FRED DATASET)

Data Processing:

Using R and relevant packages, the state names were converted to their abbreviations, and missing data were addressed.

Visualization:

Statebins were created for both rail and air GDP to visualize the economic contributions of these sectors across states. Different color palettes were used for each sector for clarity.

R Code Implementation:

```
# Load necessary libraries
library(ggplot2)
library(statebins)
library(dplyr)
```

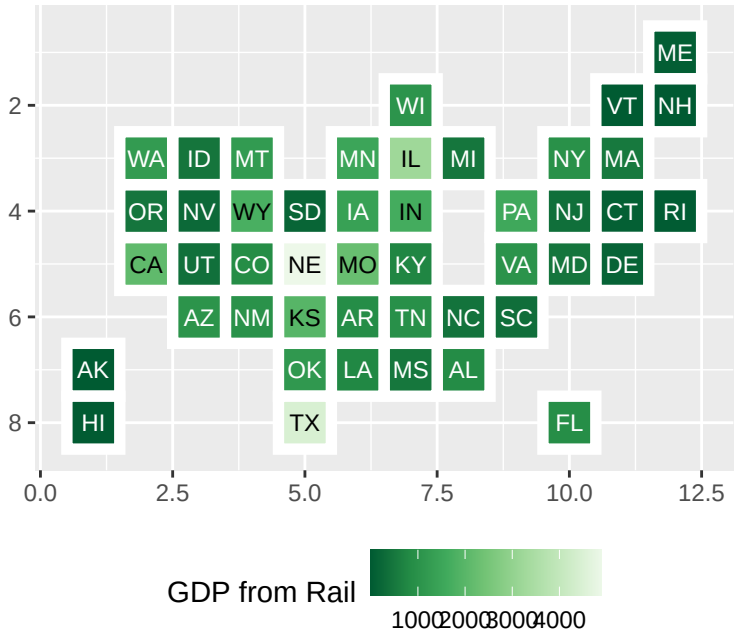
```
Attaching package: 'dplyr'

The following objects are masked from 'package:stats':

    filter, lag

The following objects are masked from 'package:base':

    intersect, setdiff, setequal, union
```

```
# Read the datasets
gdp_rail <- read.csv("/Users/naviinkumar/Desktop/gdp_raildata.csv")
gdp_air <- read.csv("/Users/naviinkumar/Desktop/gdp_airdata.csv")
# State name to abbreviation mapping
# State name to abbreviation mapping
state_map <- c("Alabama"="AL", "Alaska"="AK", "Arizona"="AZ", "Arkansas"="AR", "California
               "Colorado"="CO", "Connecticut"="CT", "Delaware"="DE", "Florida"="FL", "Geor
               "Hawaii"="HI", "Idaho"="ID", "Illinois"="IL", "Indiana"="IN", "Iowa"="IA",
               "Kansas"="KS", "Kentucky"="KY", "Louisiana"="LA", "Maine"="ME", "Maryland"=
               "Massachusetts"="MA", "Michigan"="MI", "Minnesota"="MN", "Mississippi"="MS"
               "Montana"="MT", "Nebraska"="NE", "Nevada"="NV", "New Hampshire"="NH", "New
               "New Mexico"="NM", "New York"="NY", "North Carolina"="NC", "North Dakota"="
               "Oklahoma"="OK", "Oregon"="OR", "Pennsylvania"="PA", "Rhode Island"="RI", "
               "South Dakota"="SD", "Tennessee"="TN", "Texas"="TX", "Utah"="UT", "Vermont"
               "Virginia"="VA", "Washington"="WA", "West Virginia"="WV", "Wisconsin"="WI",

# Convert full state names to abbreviations in the datasets
gdp_rail$Region <- unname(state_map[gdp_rail$Region])
gdp_air$Region <- unname(state_map[gdp_air$Region])

# Check for missing values after conversion
gdp_rail <- na.omit(gdp_rail)
```

```r
gdp_air <- na.omit(gdp_air)

# Create statebins for rail GDP
statebins_rail <- statebins(gdp_rail, state_col = "Region", value_col = "gdp_rail",
                            breaks = "quantile", labels = scales::comma) +
                  scale_fill_distiller(palette = "Greens", name = "GDP from Rail") +
                  theme(legend.position = "bottom")
```

```
Scale for fill is already present.
Adding another scale for fill, which will replace the existing scale.
```

```r
# Plot statebins for rail GDP
print(statebins_rail)
```

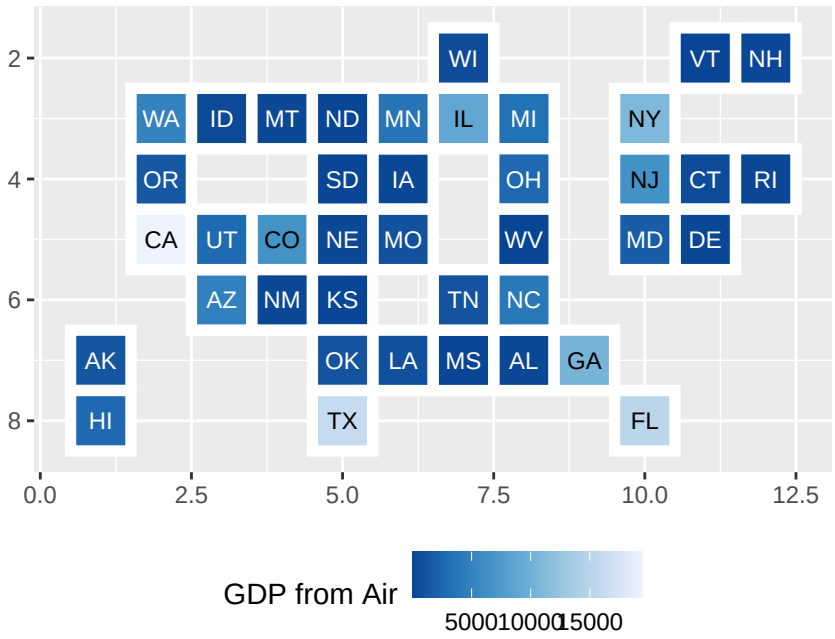



```r
# Create statebins for air GDP
statebins_air <- statebins(gdp_air, state_col = "Region", value_col = "gdp_air",
                           breaks = "quantile", labels = scales::comma) +
                 scale_fill_distiller(palette = "Blues", name = "GDP from Air") +
                 theme(legend.position = "bottom")
```

```
Scale for fill is already present.
```

Adding another scale for fill, which will replace the existing scale.

```
# Plot statebins for air GDP
print(statebins_air)
```



Analysis and Interpretation:

The statebins visualizations generated from the code above allow for a comparative analysis between the economic impacts of rail and air transport sectors in different states. The visual distinctions made by the color gradients offer insights into which states may be more dependent on these sectors.

Conclusion:

The statebins visualizations with varying shades of blue and green offer insights into the use of air and rail transportation across the United States. States with darker shades of blue show a higher usage of air transportation, indicating that these states may have more developed air transit facilities, larger airports, or a higher volume of air travel and freight. On the other hand, states with darker shades of green indicate a more significant reliance on rail transport, suggesting robust rail infrastructure that could be supporting substantial freight operations and passenger services.

These visualizations underscore the importance of transportation infrastructure in state economies. They reveal that while some states may excel in air transport, others have a comparative advantage in rail transport, reflecting the diverse economic landscapes across the country. The differences in shade intensity also highlight potential areas for infrastructure development and economic growth. It is clear from the visualizations that investing in transportation sectors where states already show strong usage can further bolster economic activity and connectivity.

In short, the states that use both rail and air transport the most are super important for making money and keeping the country running smoothly.

Future Work:

Further research could incorporate additional variables such as population, infrastructure, and investment levels to provide a more in-depth understanding of the factors influencing GDP contributions from these transport sectors.