

GR5702

Exploratory Data Analysis and Visualization

Prof. Joyce Robbins

March 23, 2017

Community Contribution idea

Columbia Statistics Club

Python Group Study Syllabus Spring 2017

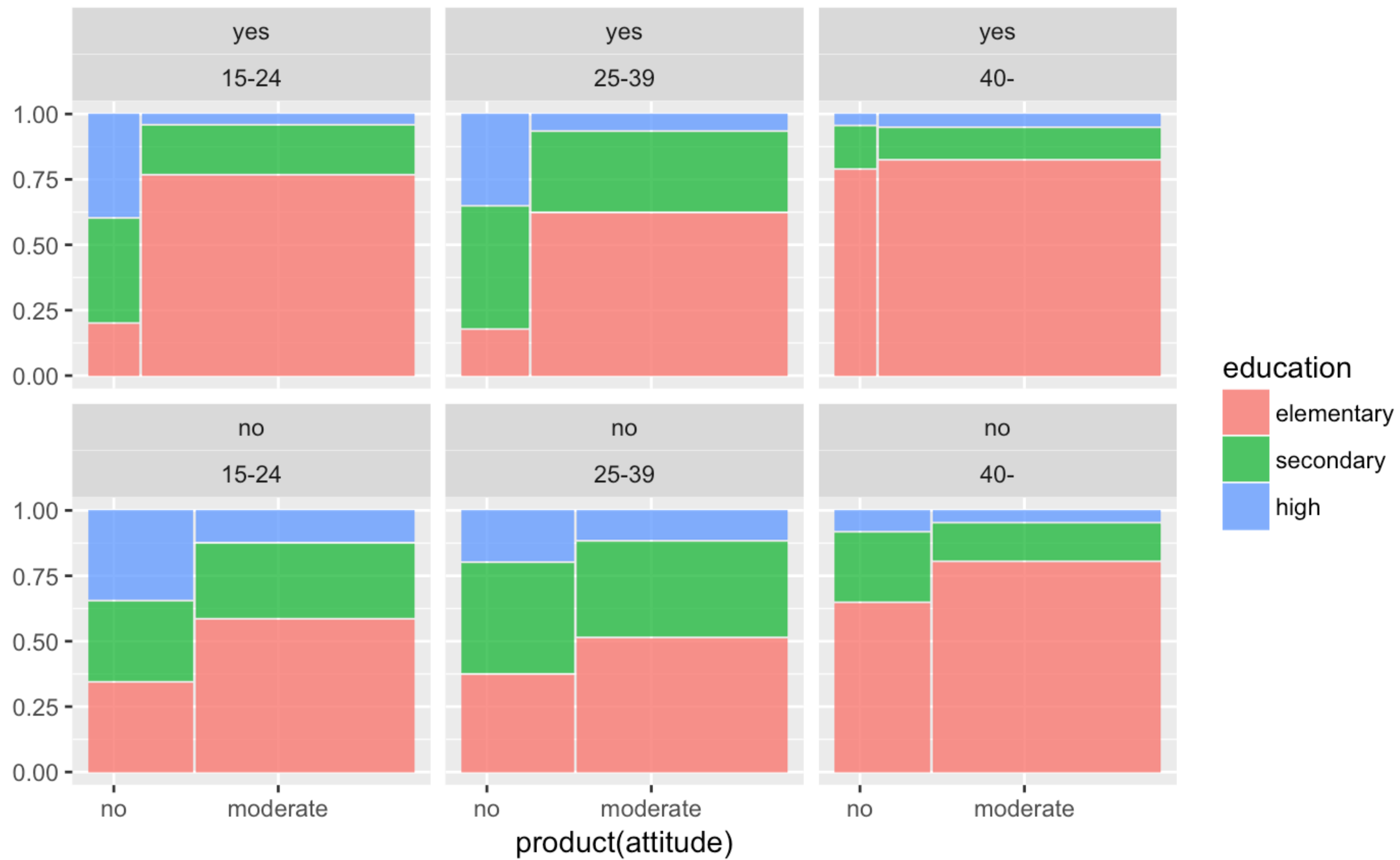
Sundays, March 26?, April 2, 9, 16, 23

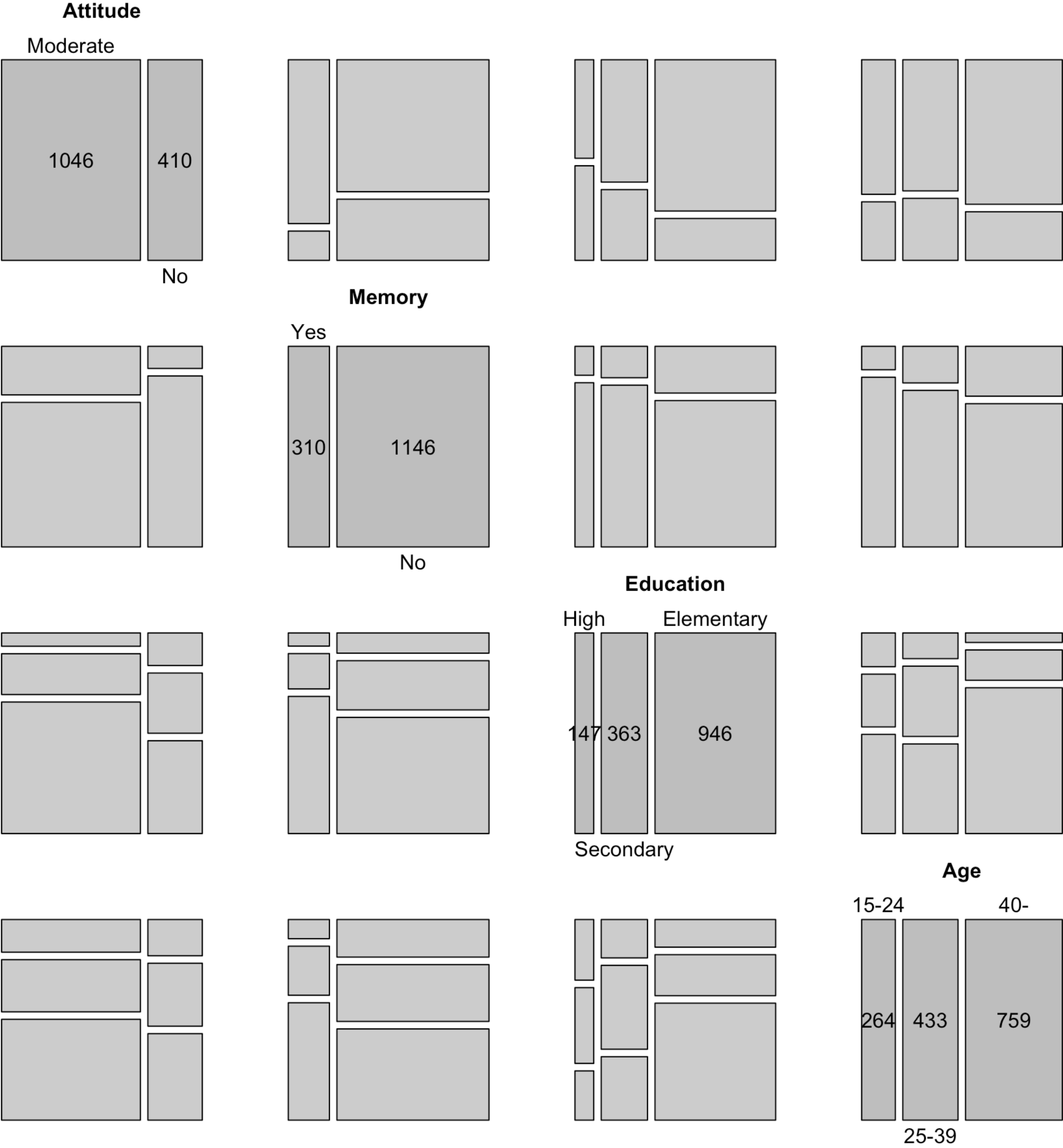
2pm - 4pm, Mathematics Room 520

Python Data Science Handbook, by Jake VanderPlas

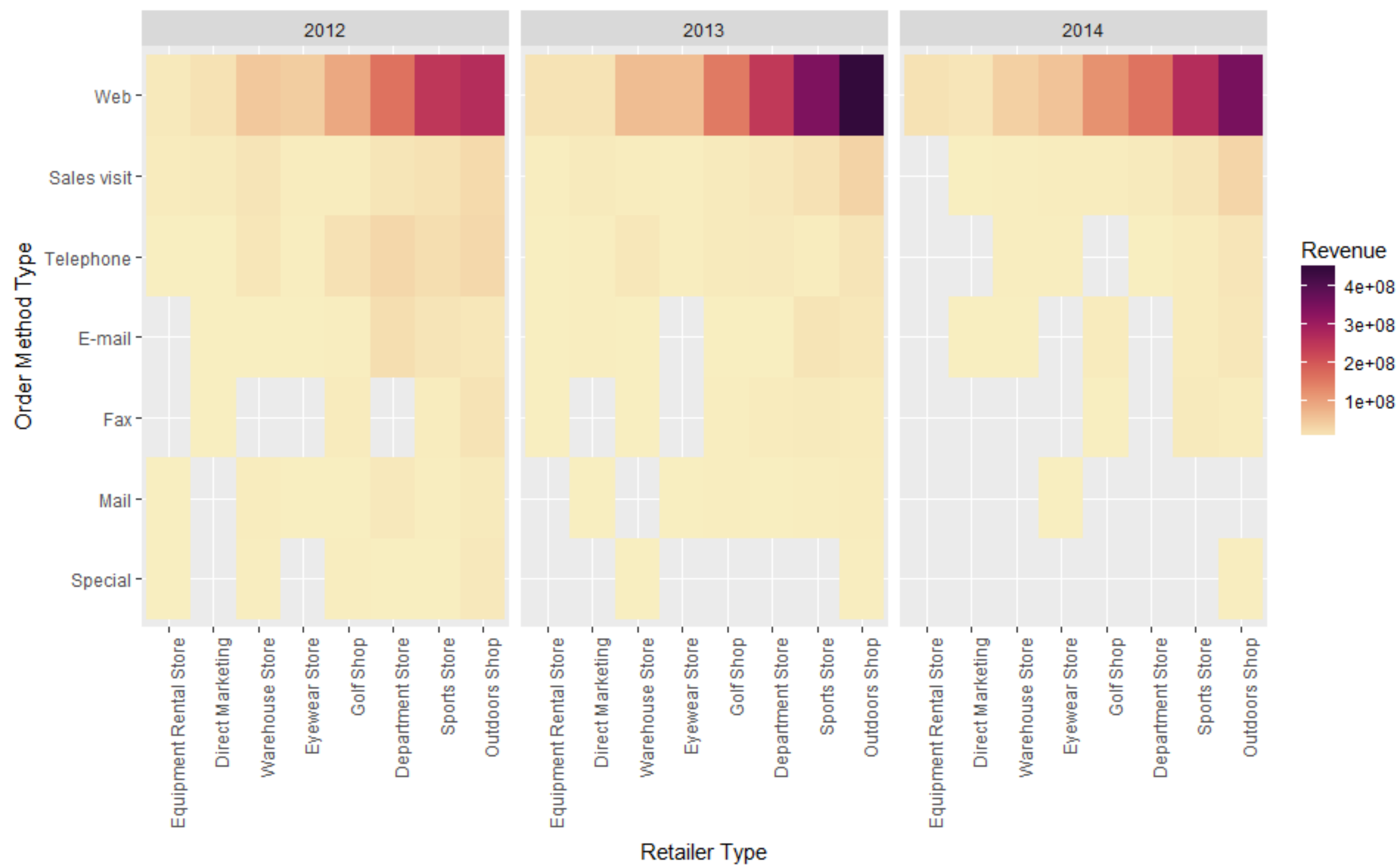
more info: [CSSPythonSyllabus.pdf](#)

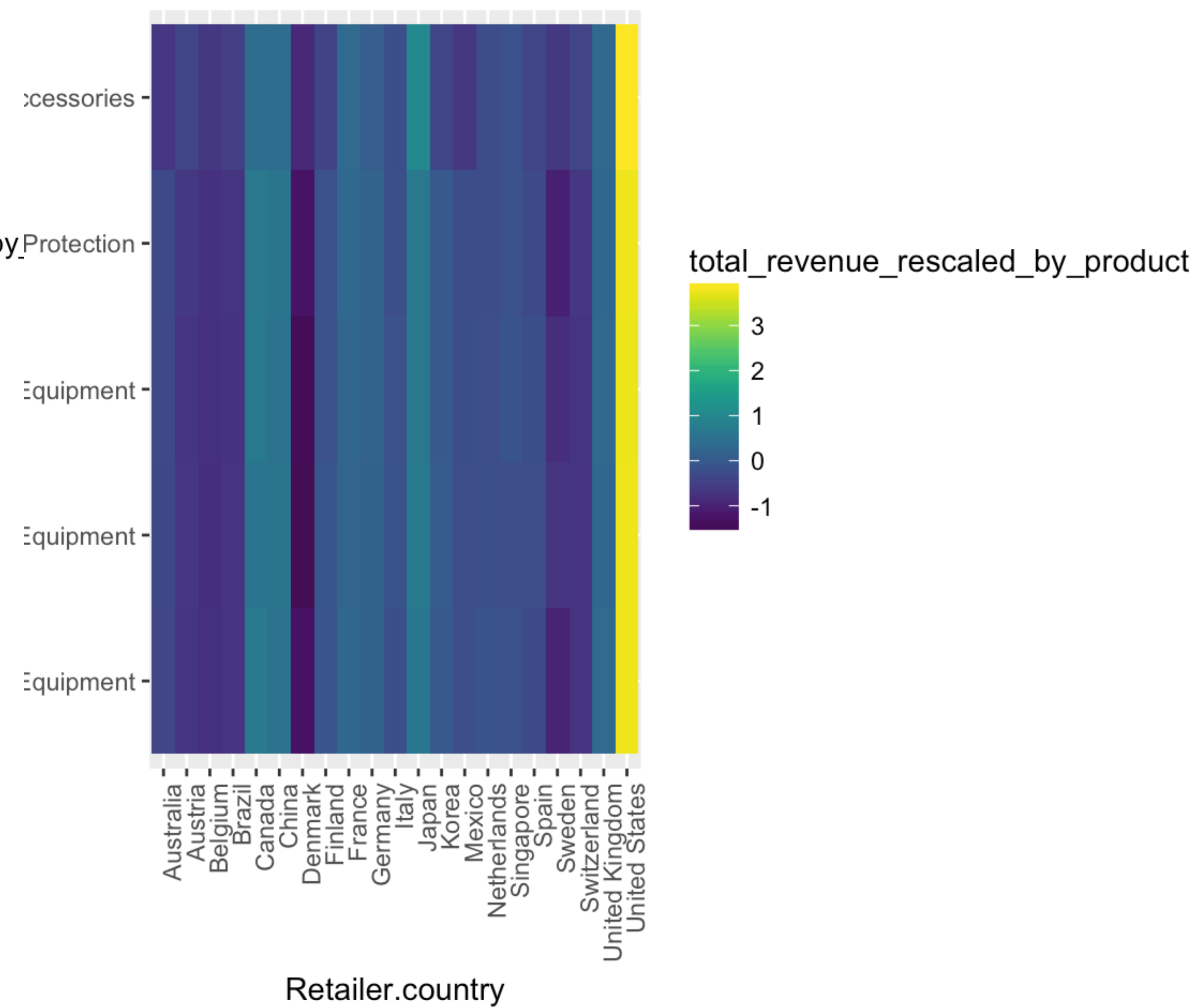
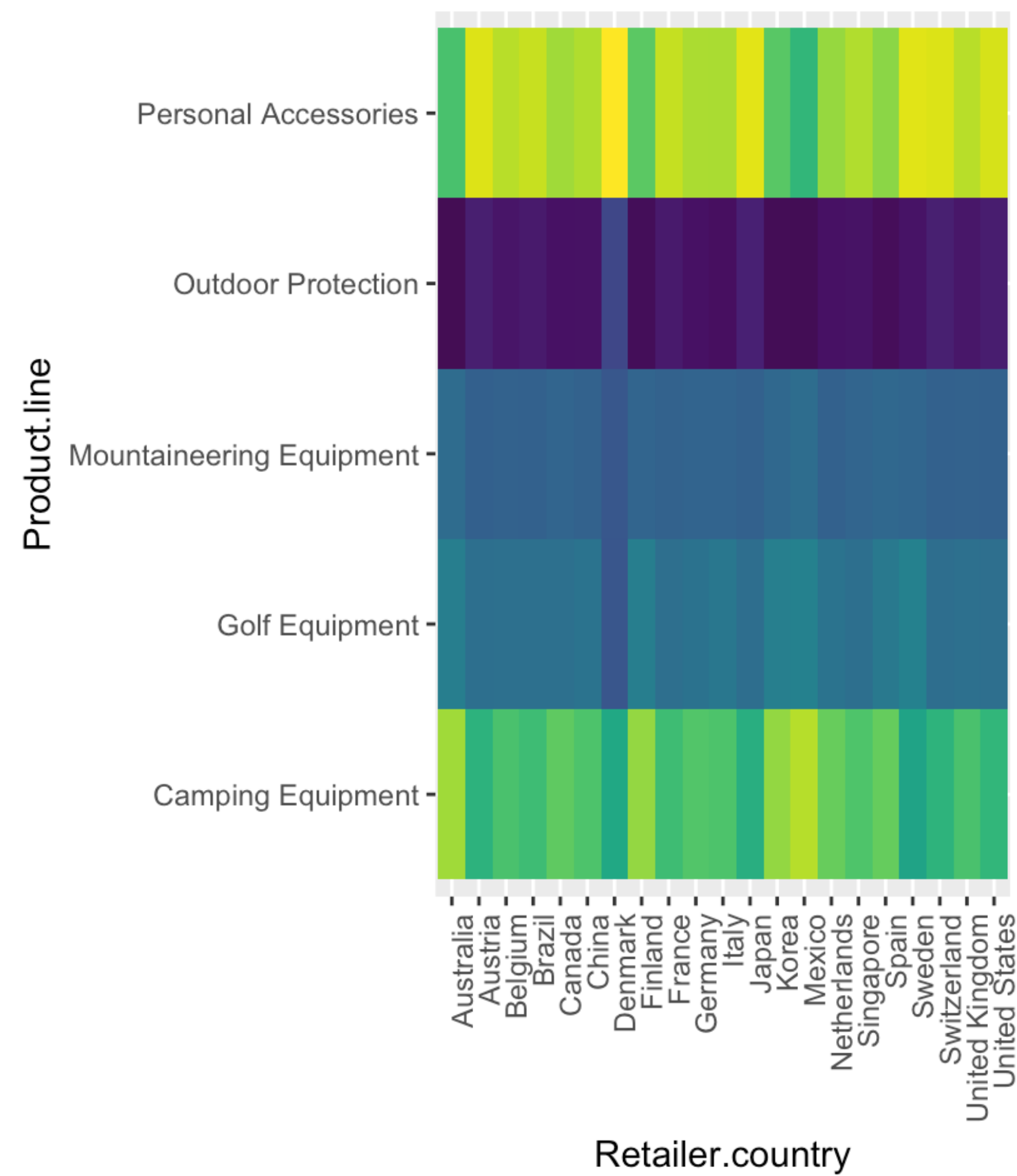
contact: Xuehan Liu (xl2615)





Plot Matrix of Barcharts and Two-Dimensional Mosaic Plots for the Punishment Dataset

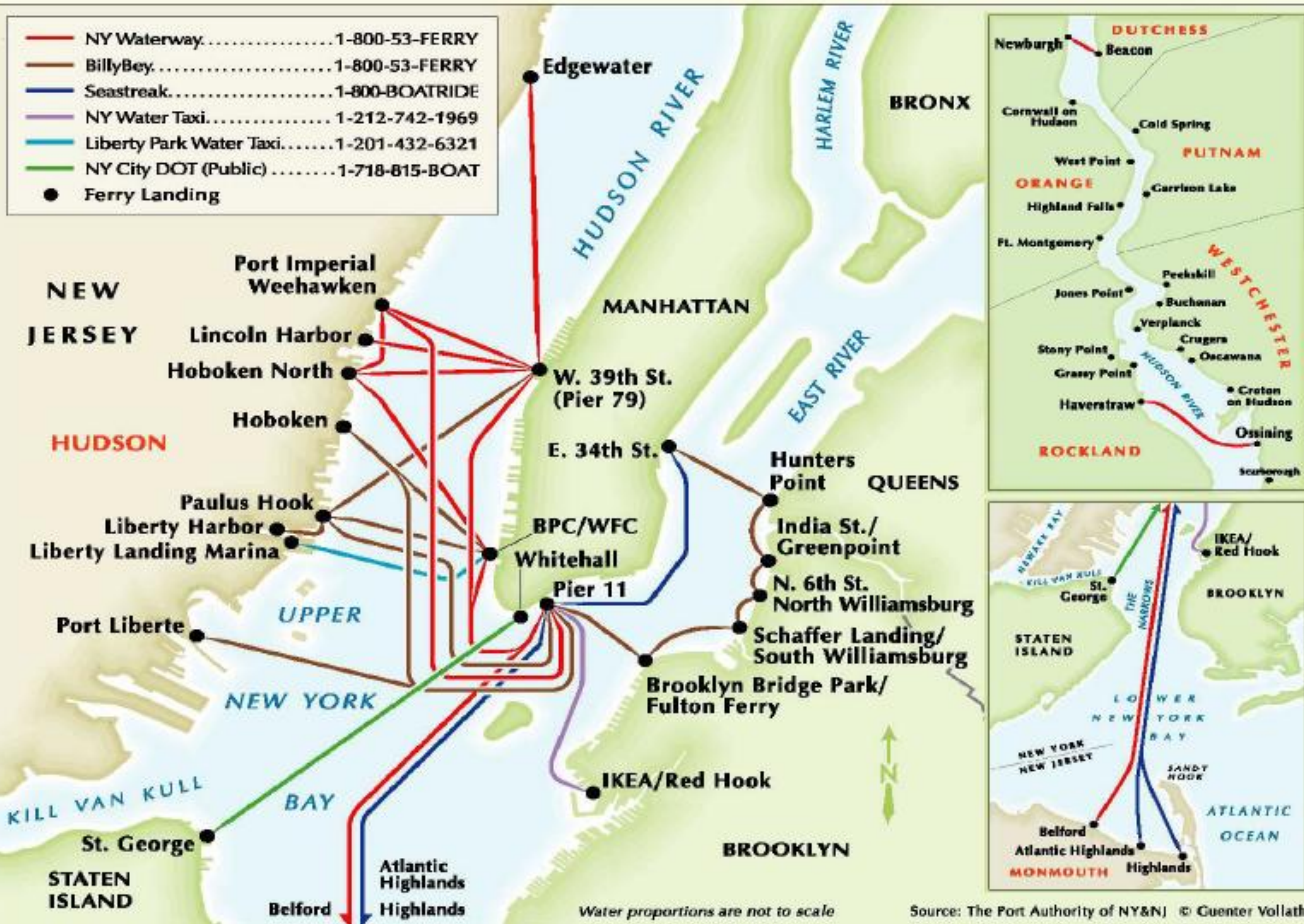




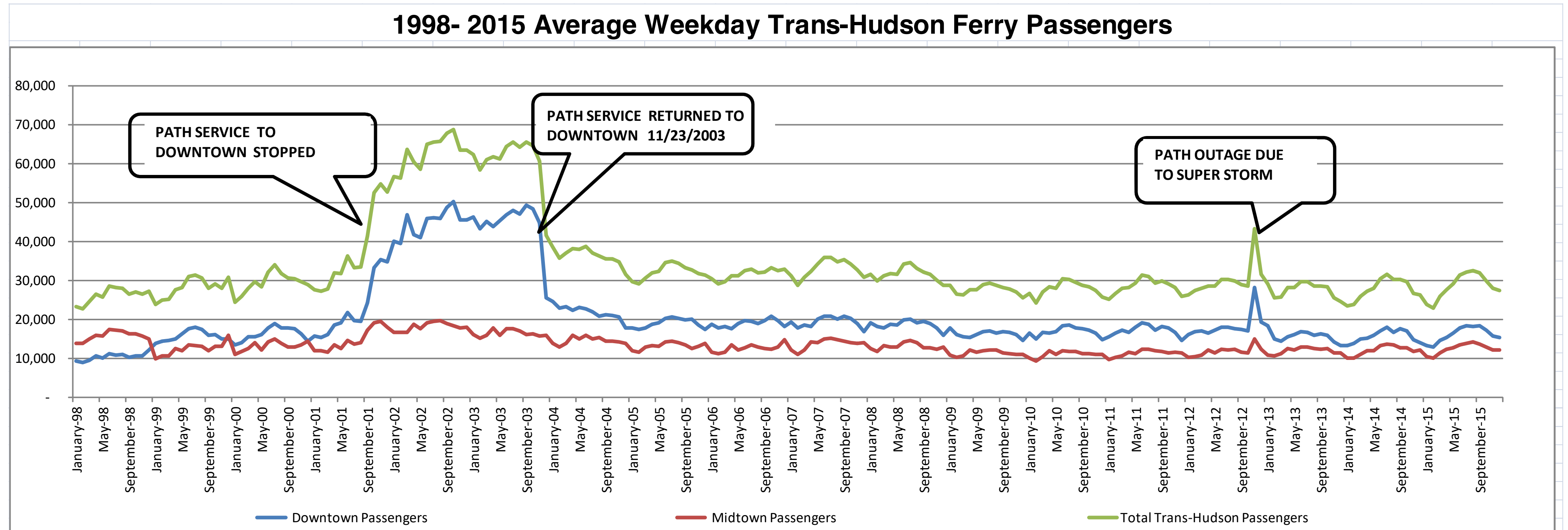
Is ferry ridership across the Hudson growing or shrinking?



NEW YORK HARBOR COMMUTER FERRY ROUTES



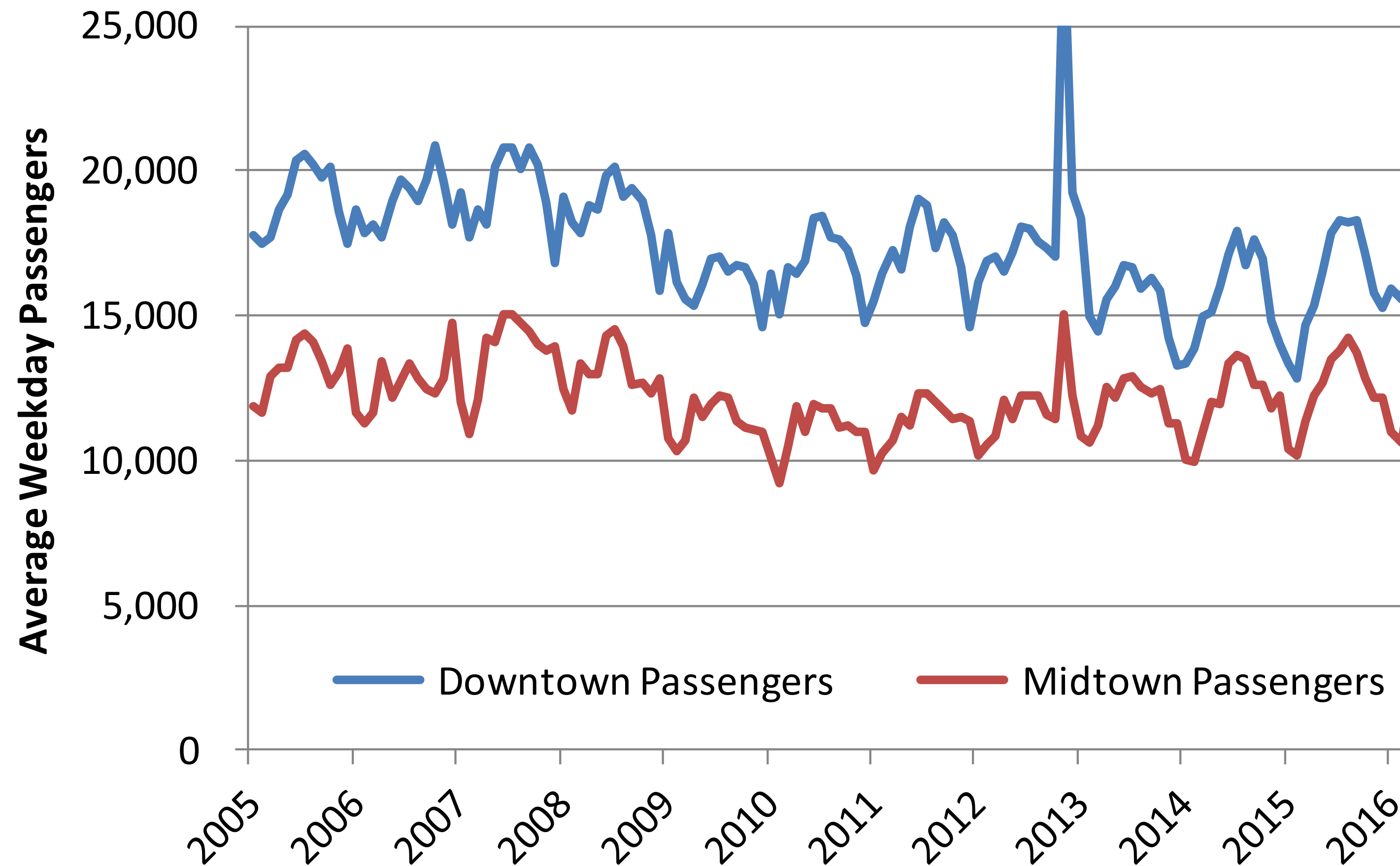
What is the trend in ridership?



Problems:

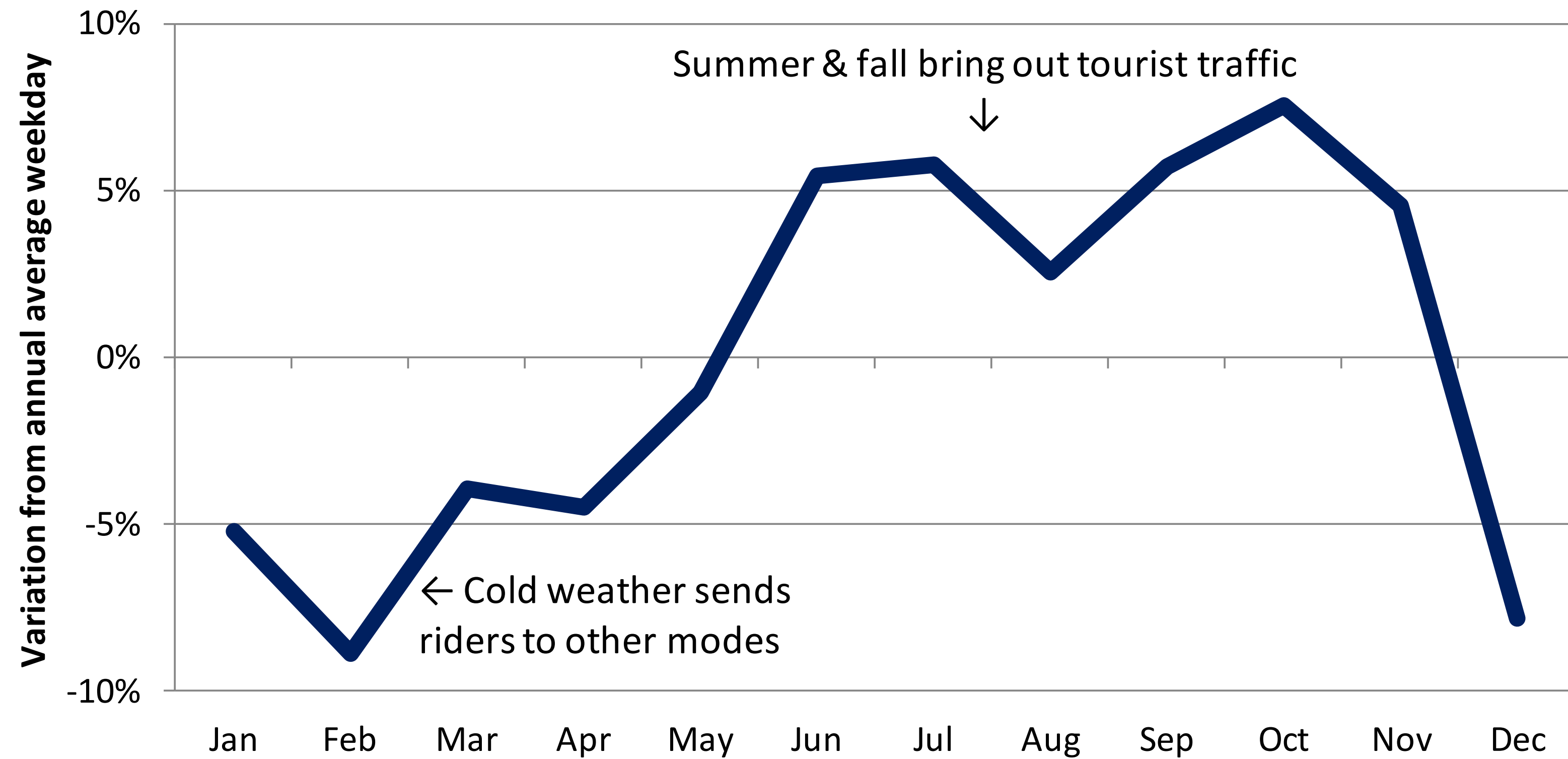
- Vertical scale makes it hard to see changes in normal ridership
- Horizontal scale shows too long a time period

Average Weekday Trans-Hudson Ferry Ridership

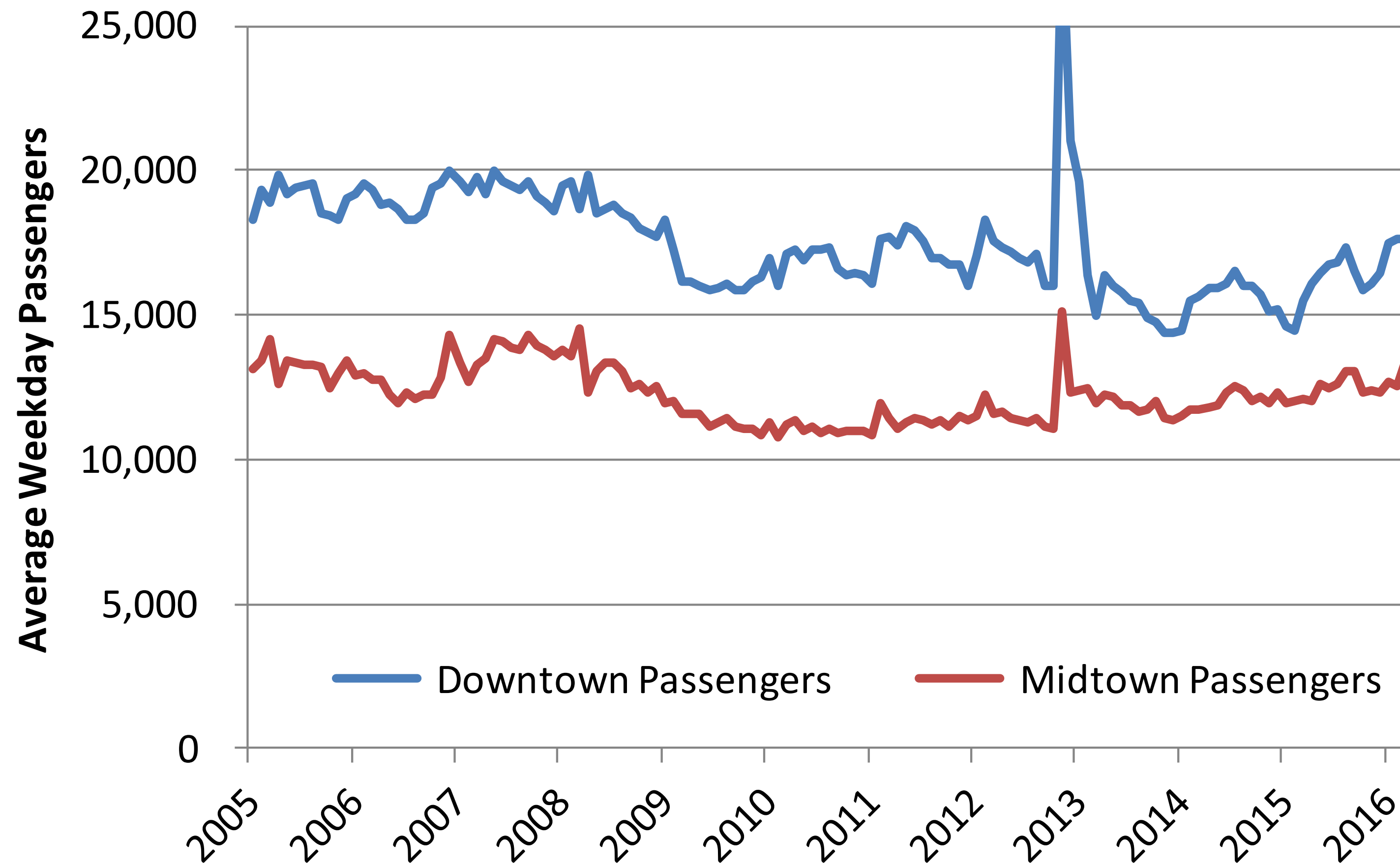


- But lots of ups and downs make it difficult to clearly see trend

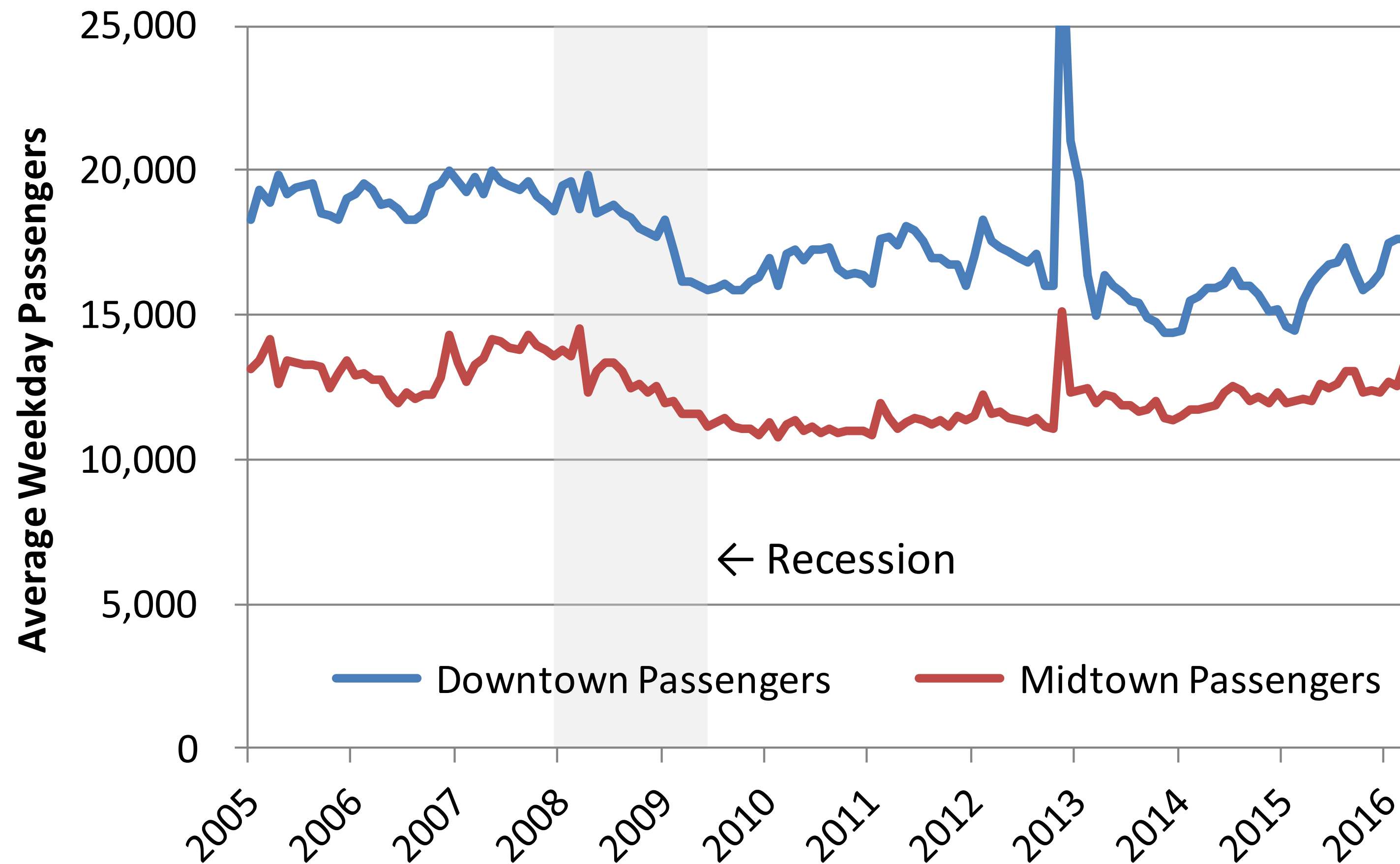
Lots of seasonal variation



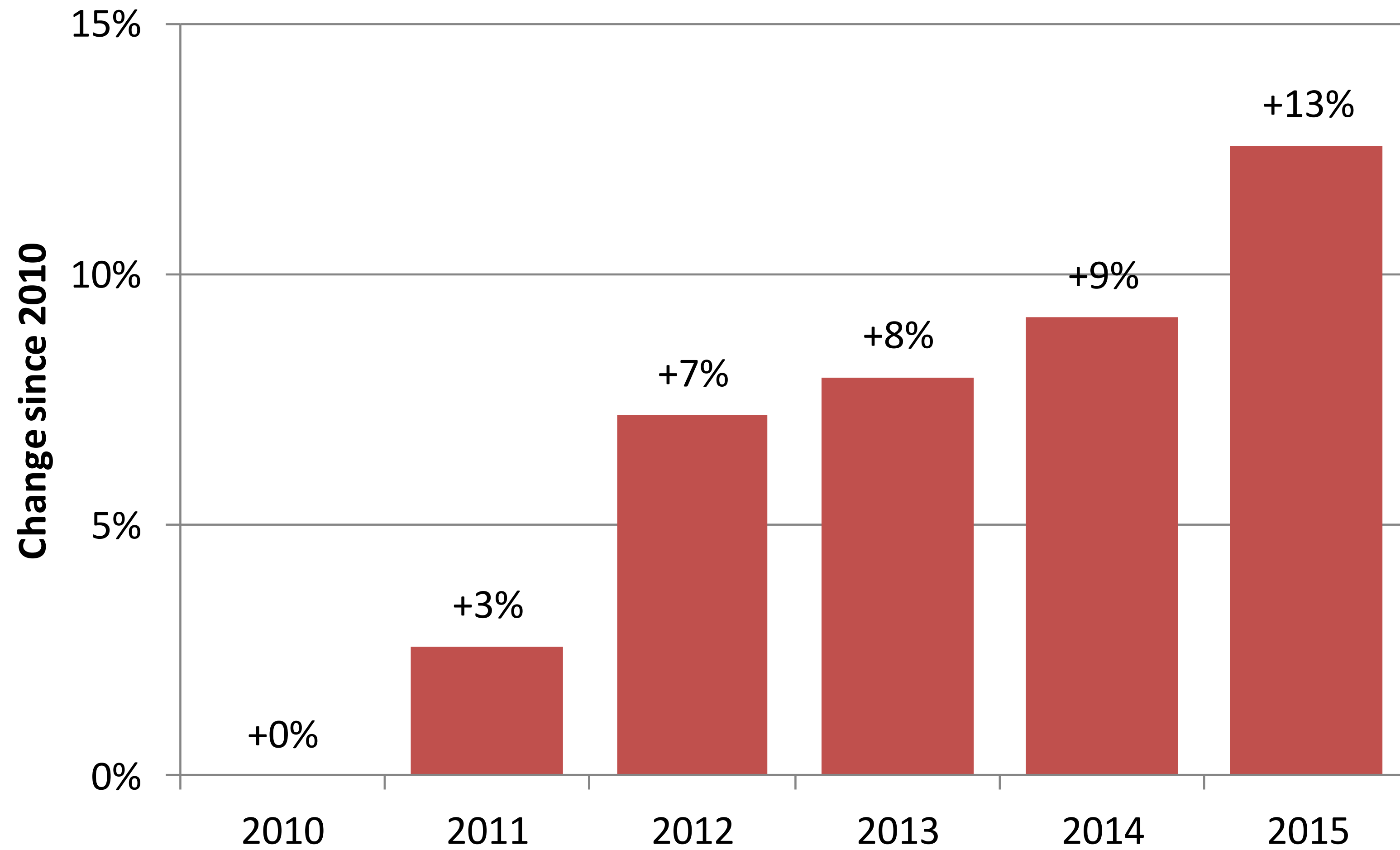
Average Weekday Trans-Hudson Ferry Ridership Seasonally Adjusted



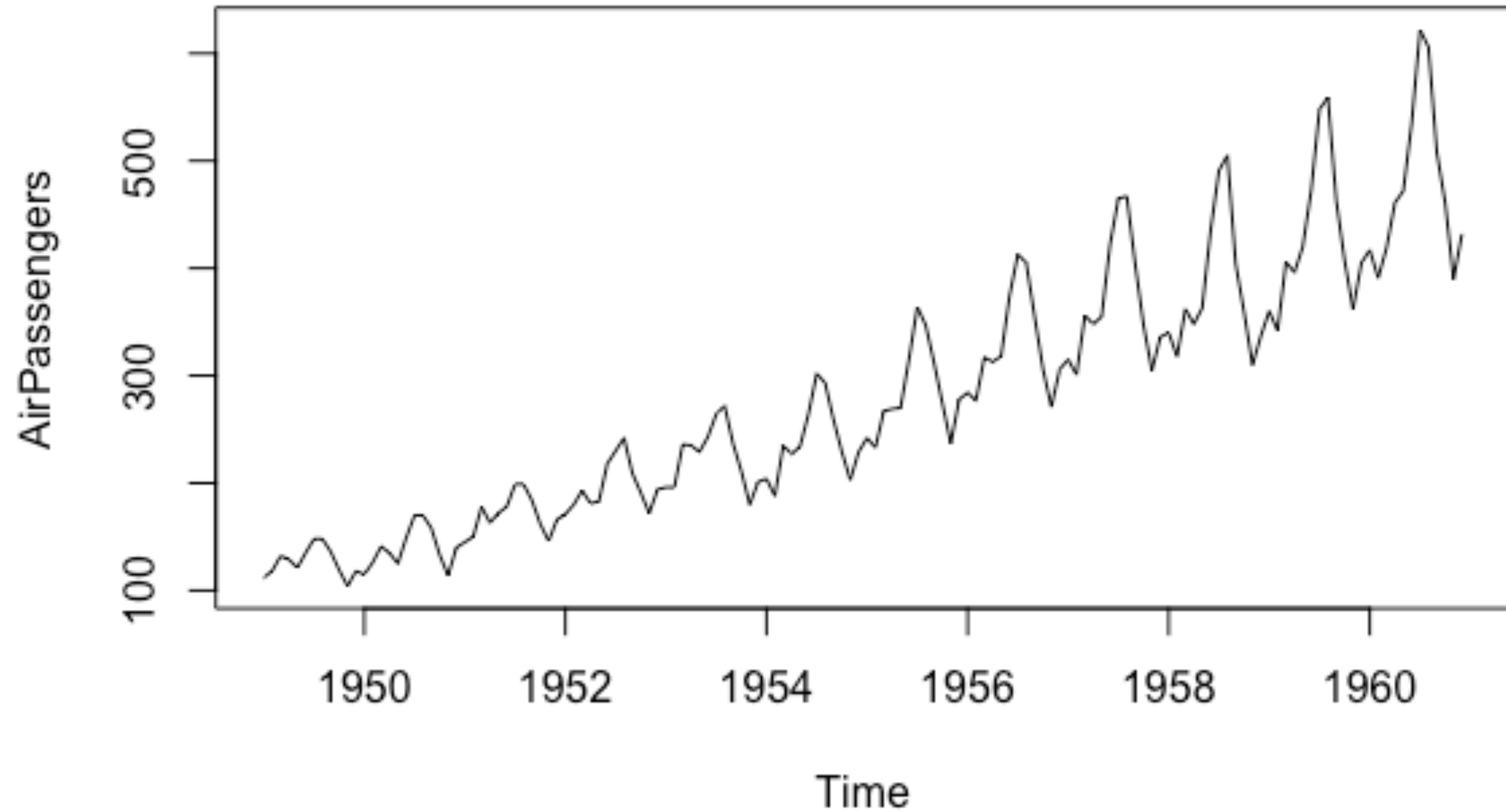
Now what can we see?



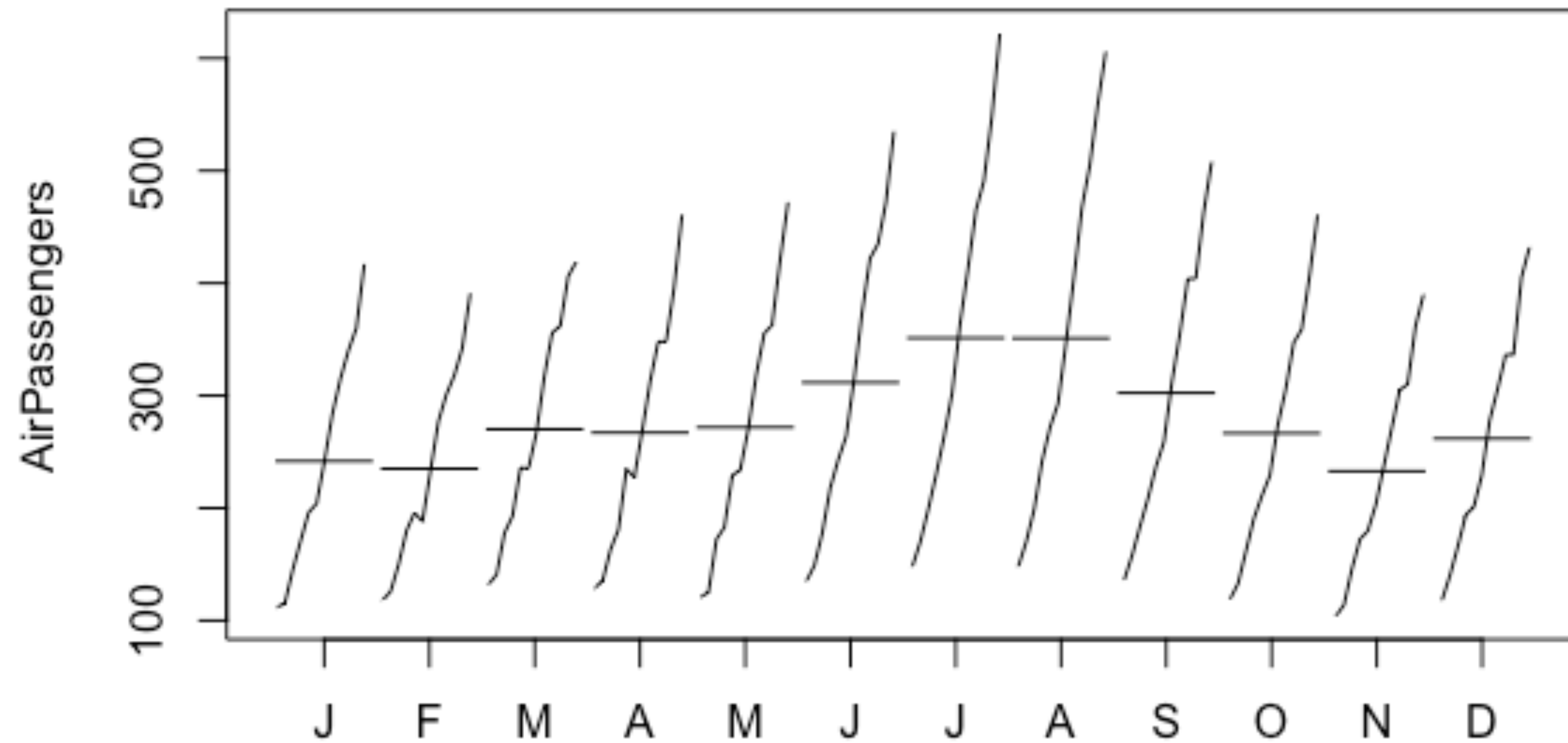
Change in Midtown ridership since 2010



```
>plot(AirPassengers)
```




```
>monthplot(AirPassengers)
```



see:

Dates.R

TimeSeries.Rmd

TimeSeries.html