



# DATA VISUALIZATION IN BUSINESS COMMUNICATION

Theory, Methods, and Tools

Gamma Iota Sigma  
2023 Regional Conference  
Fort Worth, TX

# INTRODUCTIONS



**Dalesa Bady, ACAS, MAAA**  
*Actuary, GuideOne Insurance*

Associate of the Casualty Actuarial Society (CAS)  
12 years of experience in the property and casualty insurance industry.  
Fav areas of work: ratemaking, product development, and predictive analytics.  
Passionate about coaching and development.  
CAS University Liaison for UT Austin/UT Dallas and Chair of IABA Scholarship Committee  
Likes to travel and play flag football.



**Bryce Chamberlain, ASA, MScA**  
*Principal, Oliver Wyman Actuarial Consulting*

Associate of the Society of Actuaries (SOA)  
Master of Science in Analytics, University of Chicago  
Leads a team at building business intelligence apps for the web using R Shiny.  
Passionate about data visualization, user-friendly design, and efficiency.  
Lead developer for R packages: easyr, hcslim.  
Likes to play video games and meet friends at cocktail bars.

# AGENDA



**1**

Why Visualize?

**2**

How We Think Visually

**3**

Things To Avoid

**4**

Problems & Tools to Solve Them

The last slide will show a link to this deck + all the resources mentioned.

# WHY VISUALIZE?

“

How many of you pay attention to road signs  
while driving?

How much harder would it be if road signs were a full body of text?

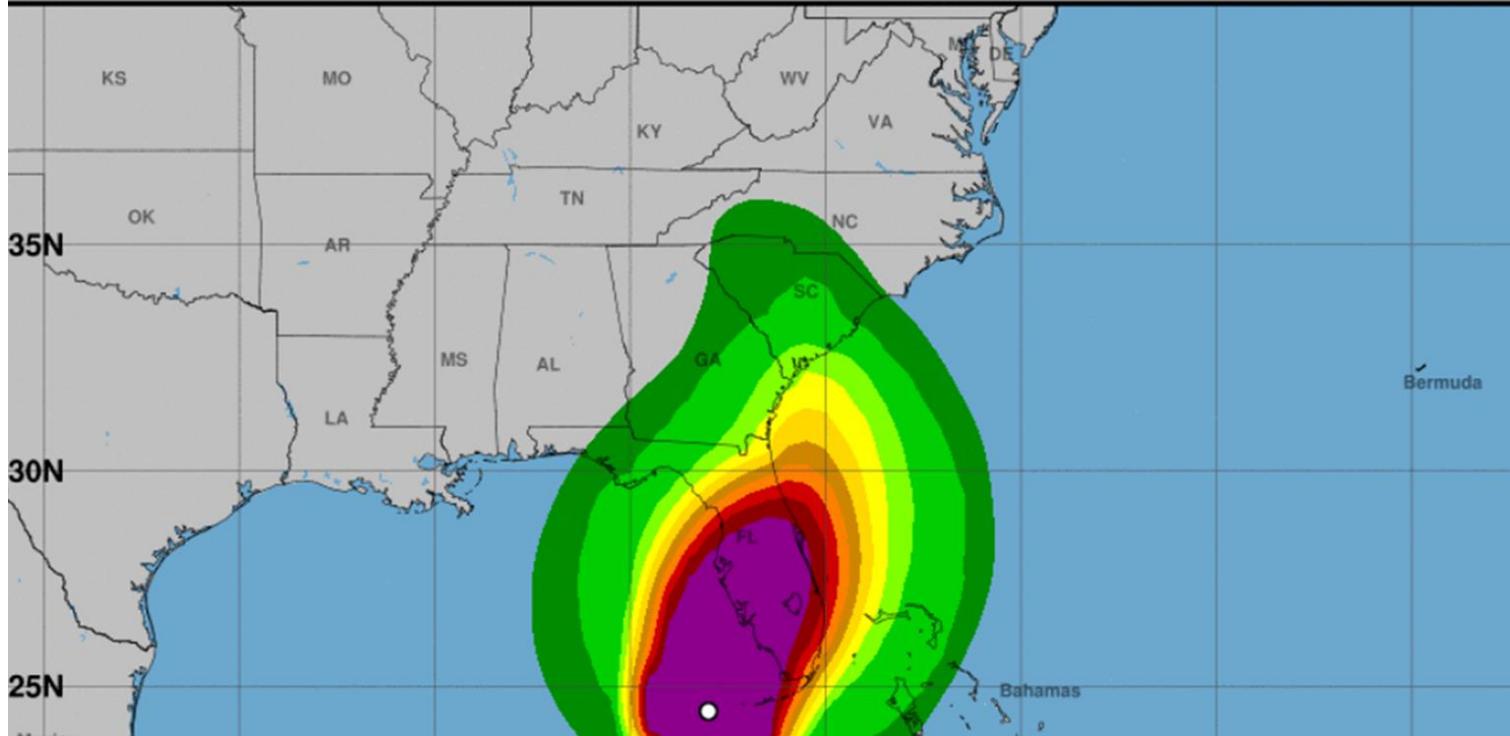


Beware of the road ahead. It might have rained or snowed and it could be slippery, which might cause an incident. Be careful.



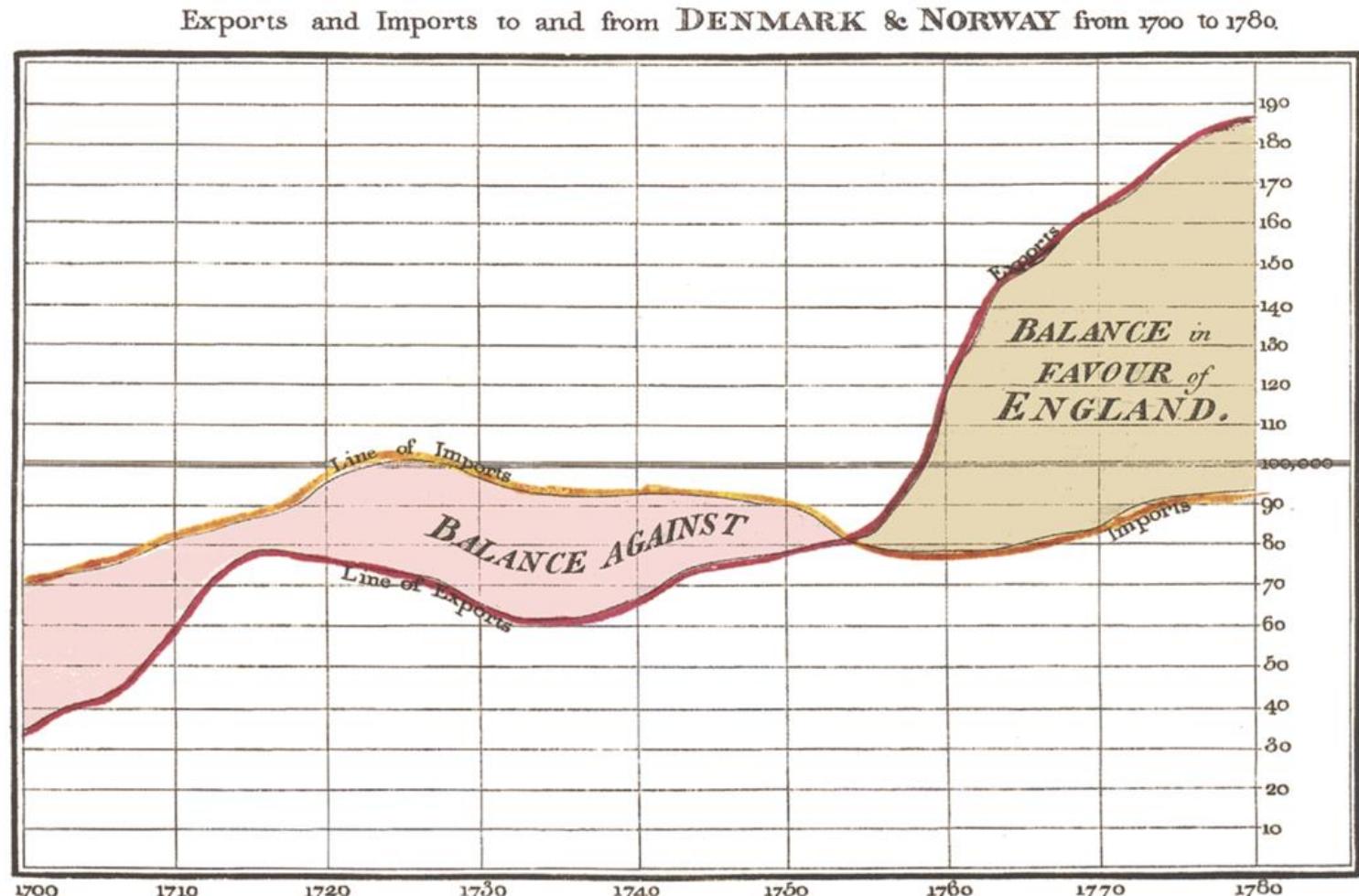
## Tropical-Storm-Force Wind Speed Probabilities (Preliminary)

For the 120 hours (5.0 days) from 8 AM EDT WED SEP 28 to 8 AM EDT MON OCT 03



Would this message be as powerful with a table of numbers?

“The Commercial and  
Political Atlas”  
- William Playfair (1786)



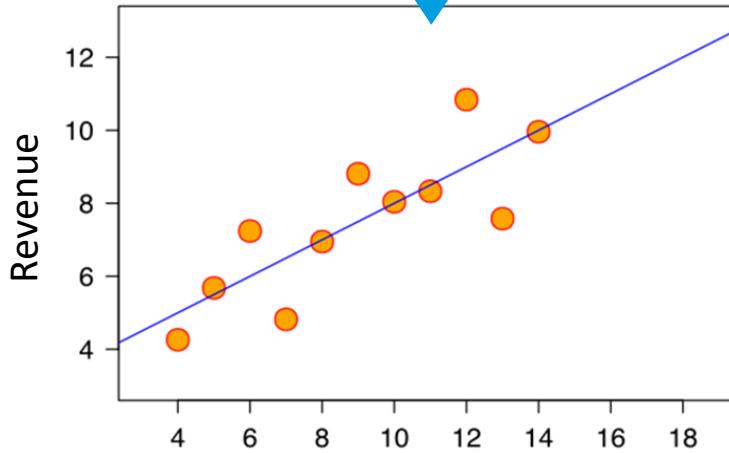
“

Well, just show me the numbers.

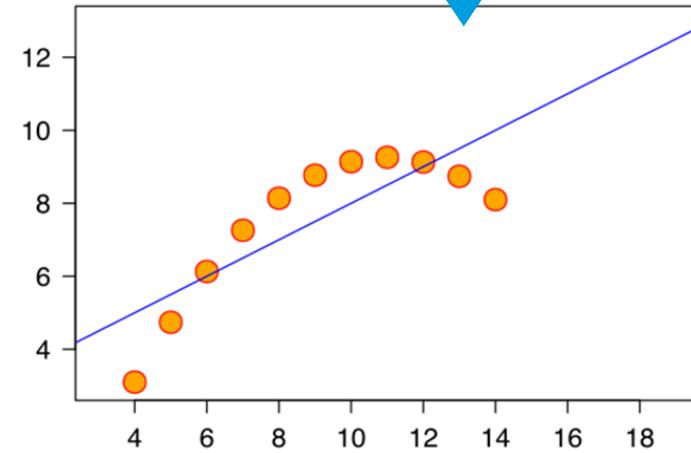
Let's take a look at an example.

| Agent | Mean Claims (\$millions) | Mean Revenue (\$millions) | Correlation |
|-------|--------------------------|---------------------------|-------------|
| A     | 9                        | 7.5                       | 0.816       |
| B     | 9                        | 7.5                       | 0.816       |
| C     | 9                        | 7.5                       | 0.816       |
| D     | 9                        | 7.5                       | 0.816       |

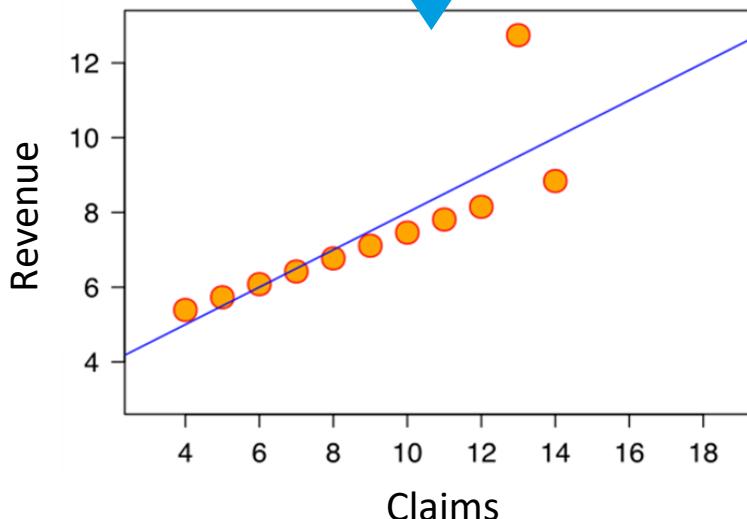
Revenue increases with claims.



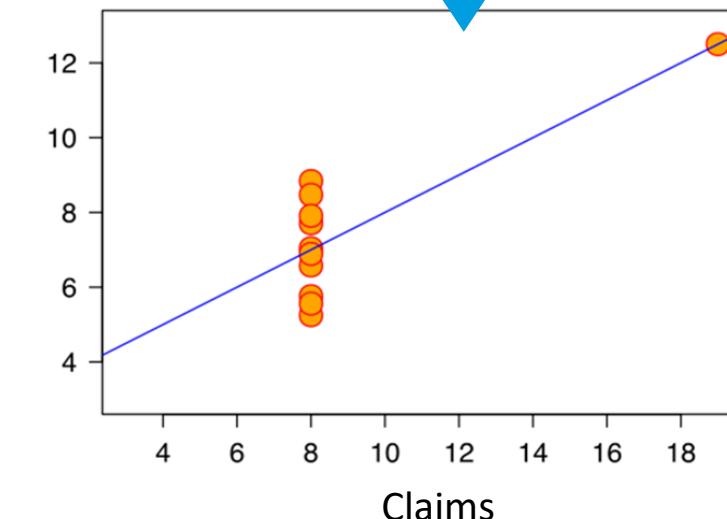
Revenue capped.



Revenue outlier.



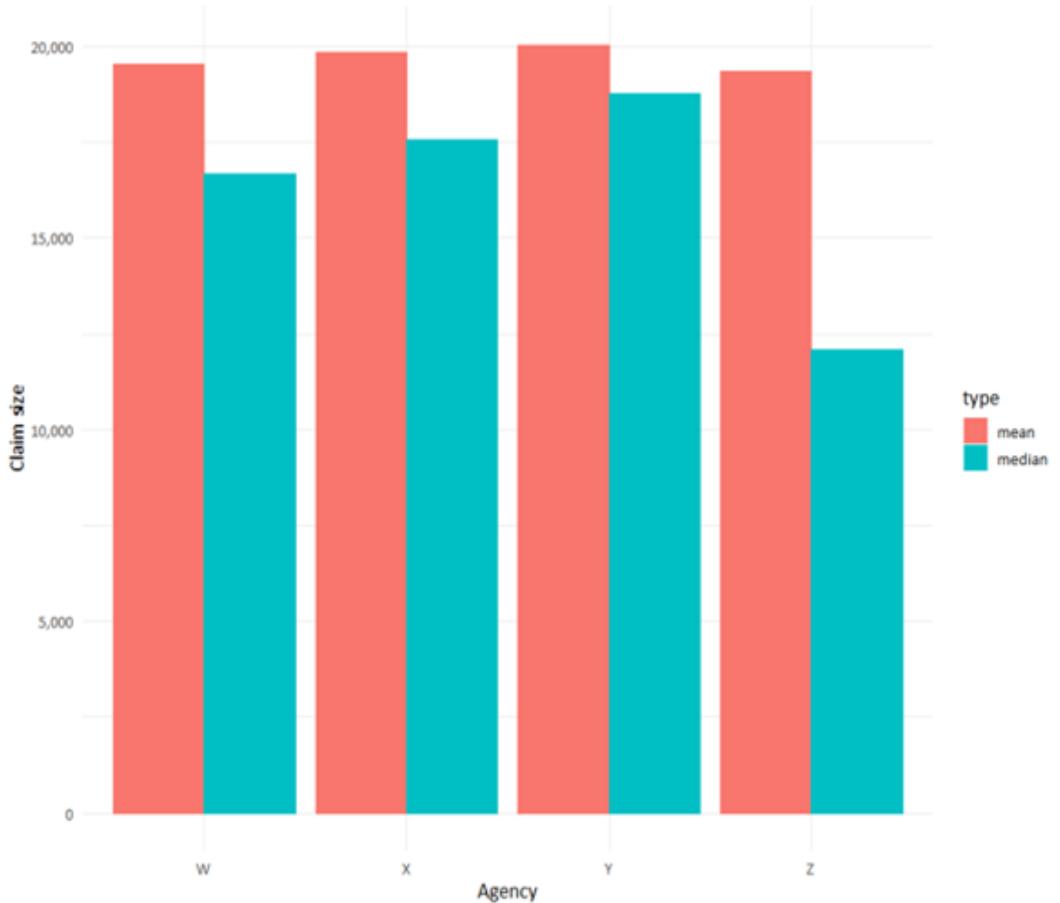
Revenue and amount outlier.



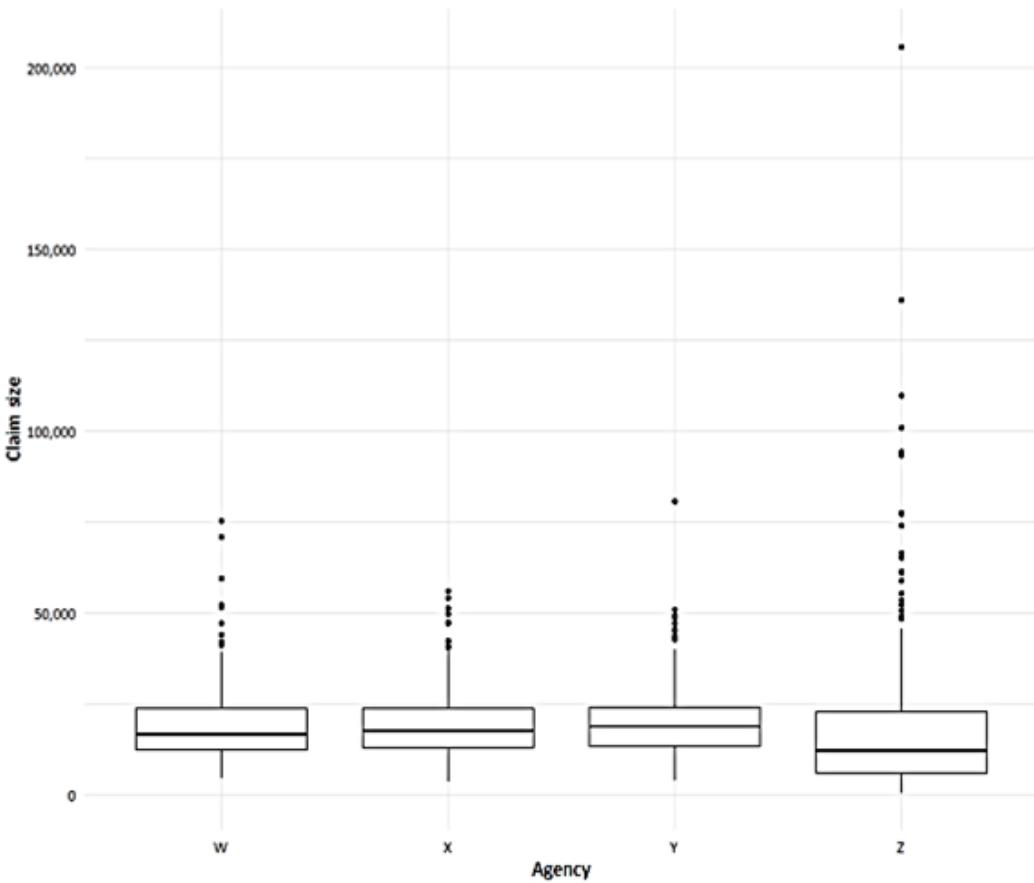
“

Visualization is useful when communicating  
to a non-technical person or audience.

| Agency | Mean<br>Claim Size | Median<br>Claim Size |
|--------|--------------------|----------------------|
| W      | 19,533             | 16,675               |
| X      | 19,829             | 17,577               |
| Y      | 20,039             | 18,758               |
| Z      | 19,363             | 12,097               |



| Agency | Mean<br>Claim Size | Median<br>Claim Size |
|--------|--------------------|----------------------|
| W      | 19,533             | 16,675               |
| X      | 19,829             | 17,577               |
| Y      | 20,039             | 18,758               |
| Z      | 19,363             | 12,097               |



# **HOW DO WE THINK VISUALLY?**

**99999999**

**9999999999**

**999,999,999**

**99,999,999,999**

**999,999,999**

**99,999,999,999**

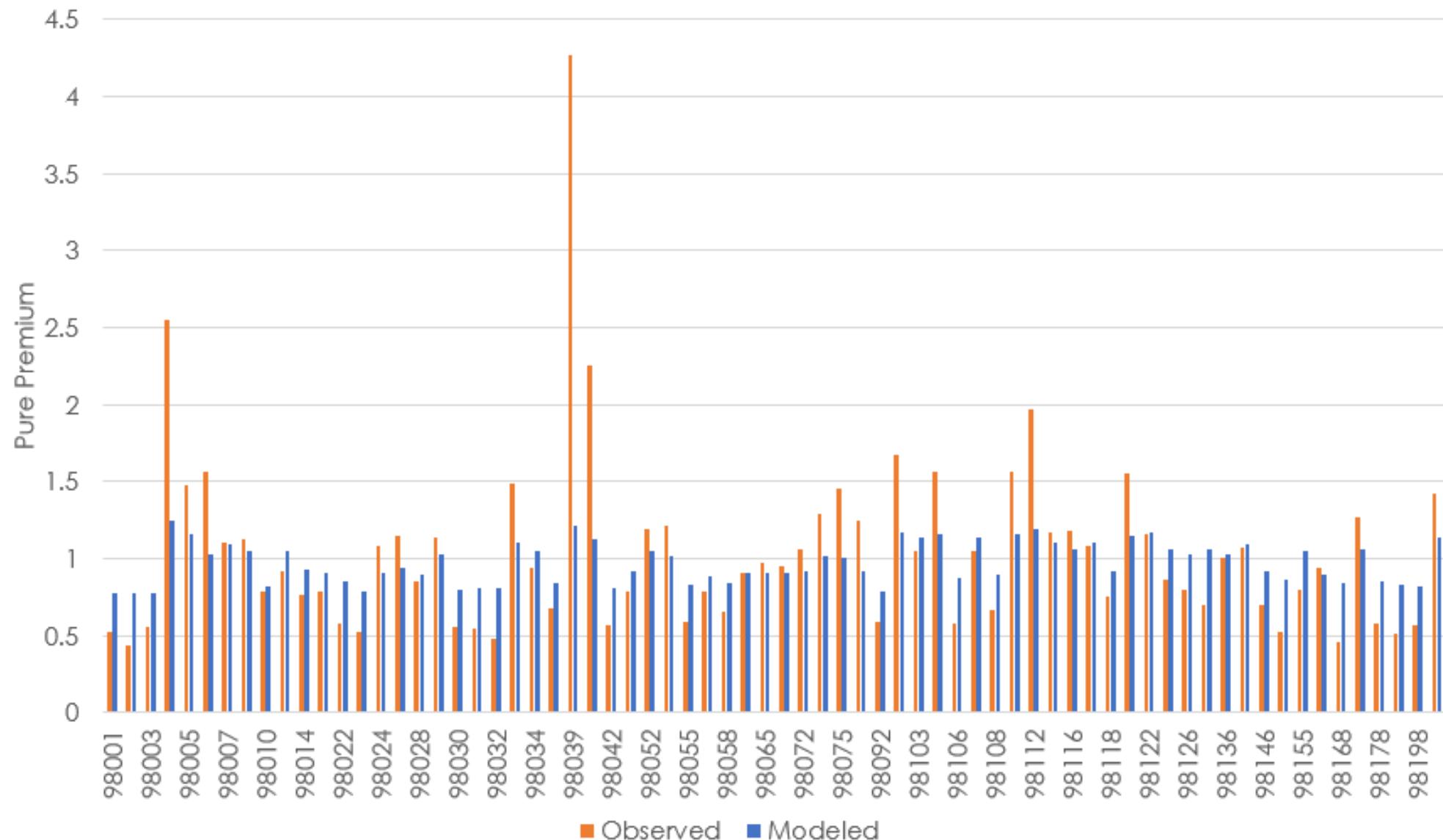


Mapping the values in a visual space helps us understand the enormity of the difference!

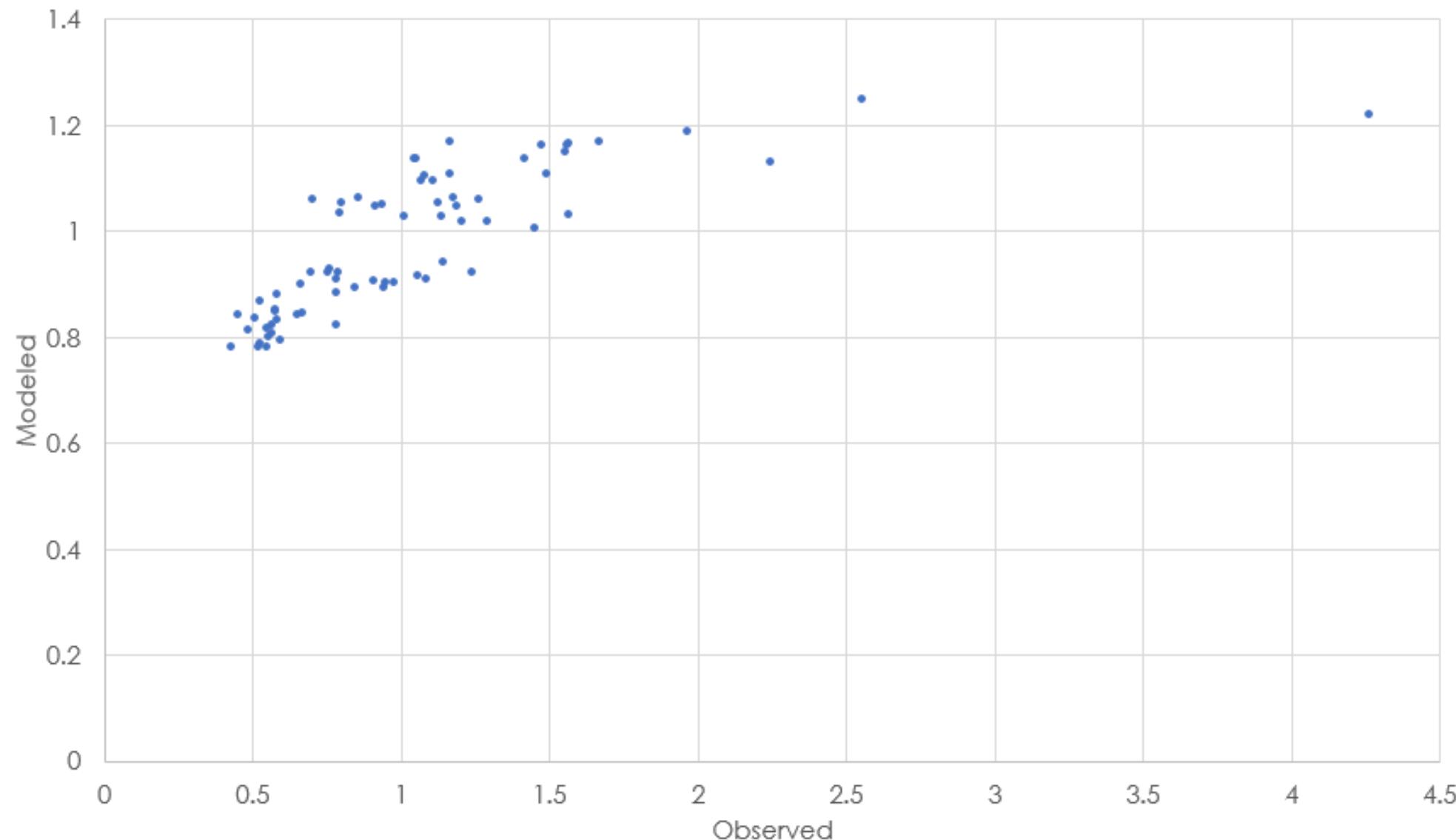
“

Let's take a look at an example of an insurance dataset visualized in two ways.

### Observed vs. Modeled Pure Premium by ZIP Code

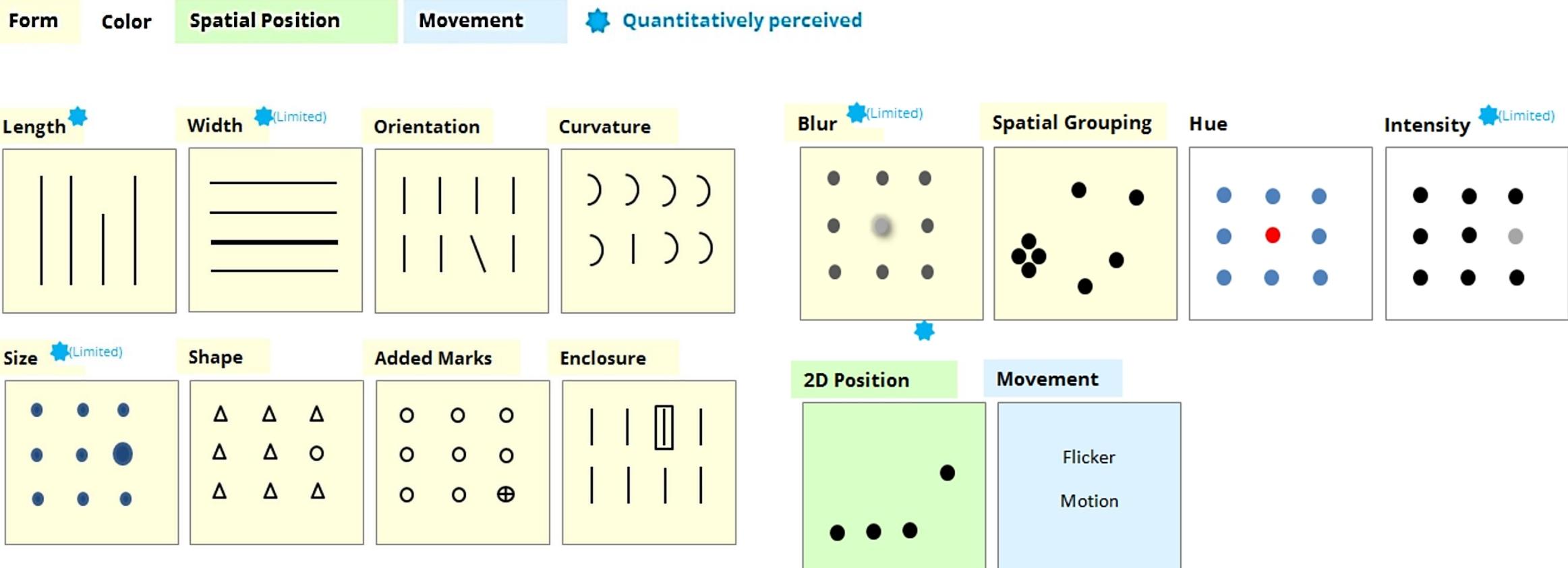


Observed vs. Modeled Pure Premium by ZIP Code

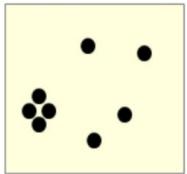


# PRE-ATTENTIVE ATTRIBUTES

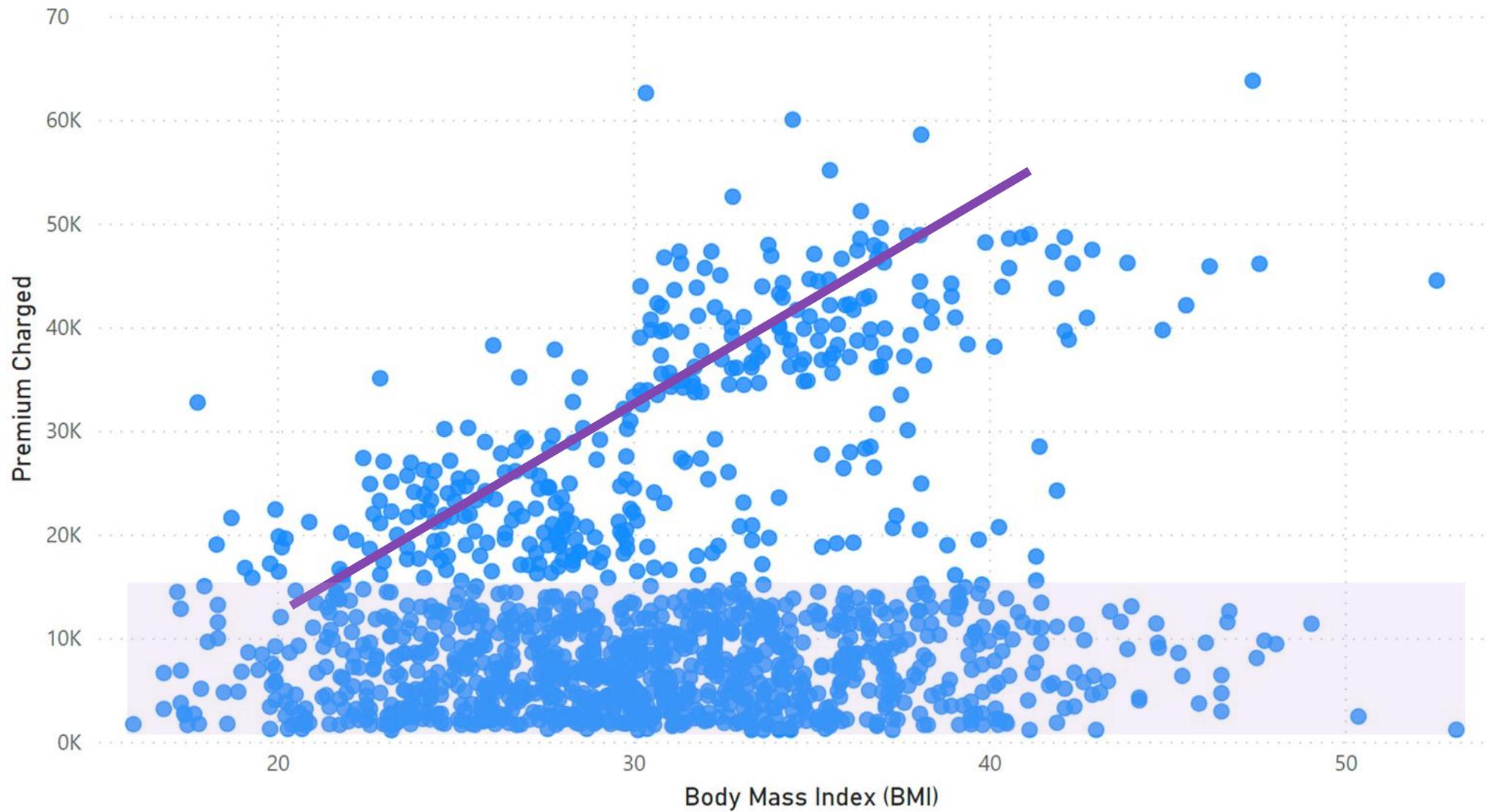
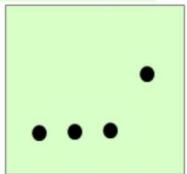
# TOOLS FOR COMMUNICATING VISUALLY

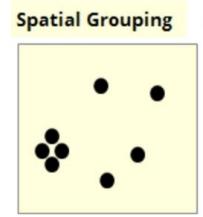


Spatial Grouping

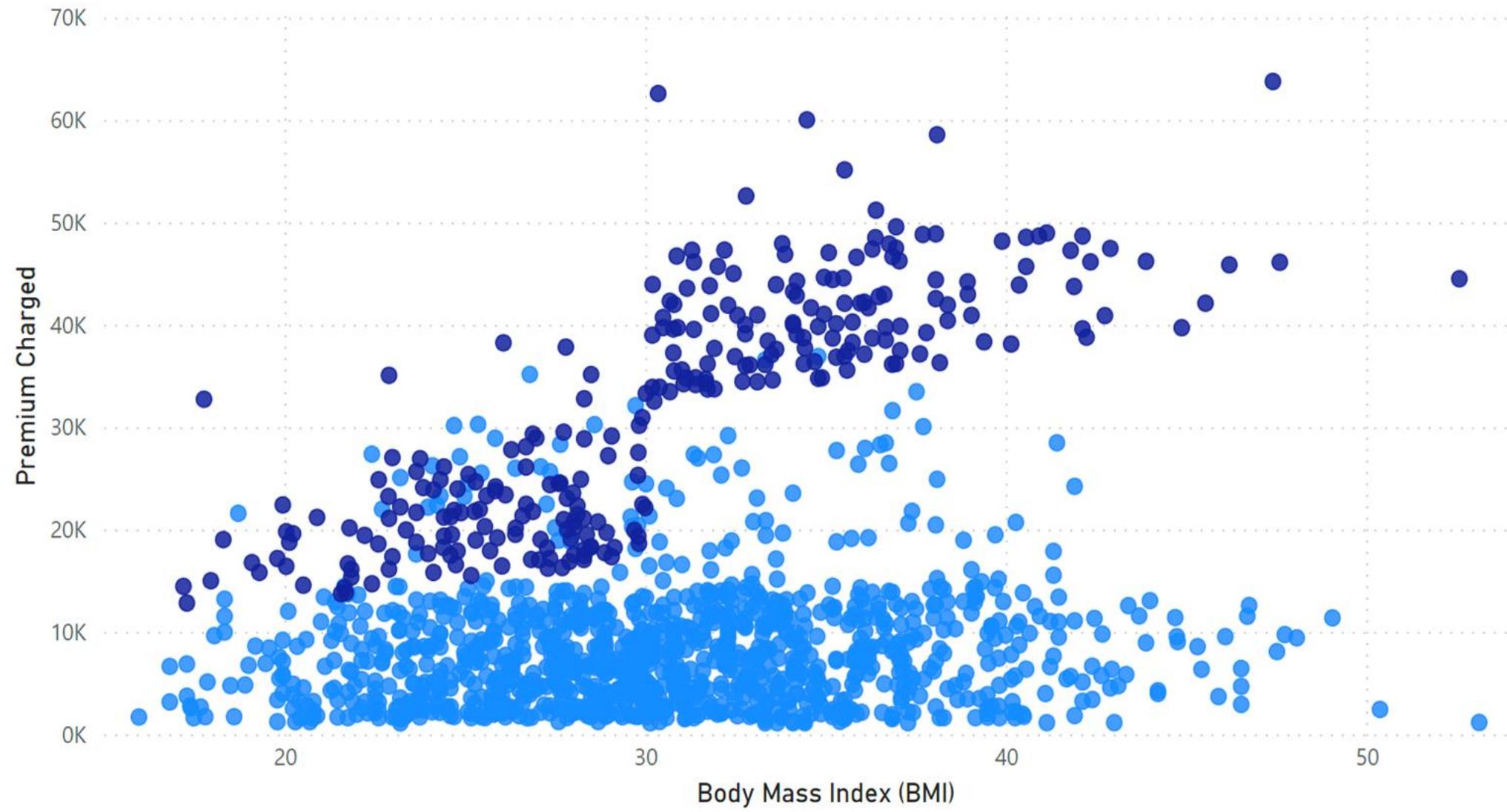
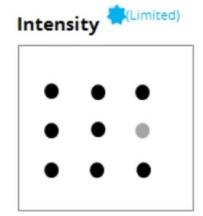
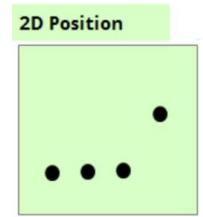


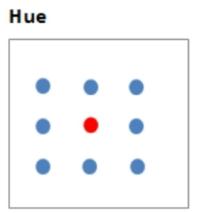
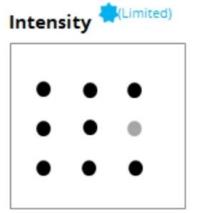
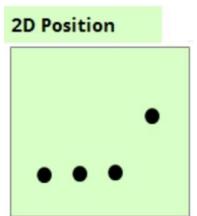
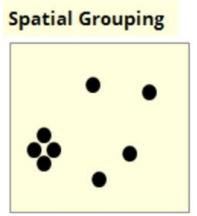
2D Position



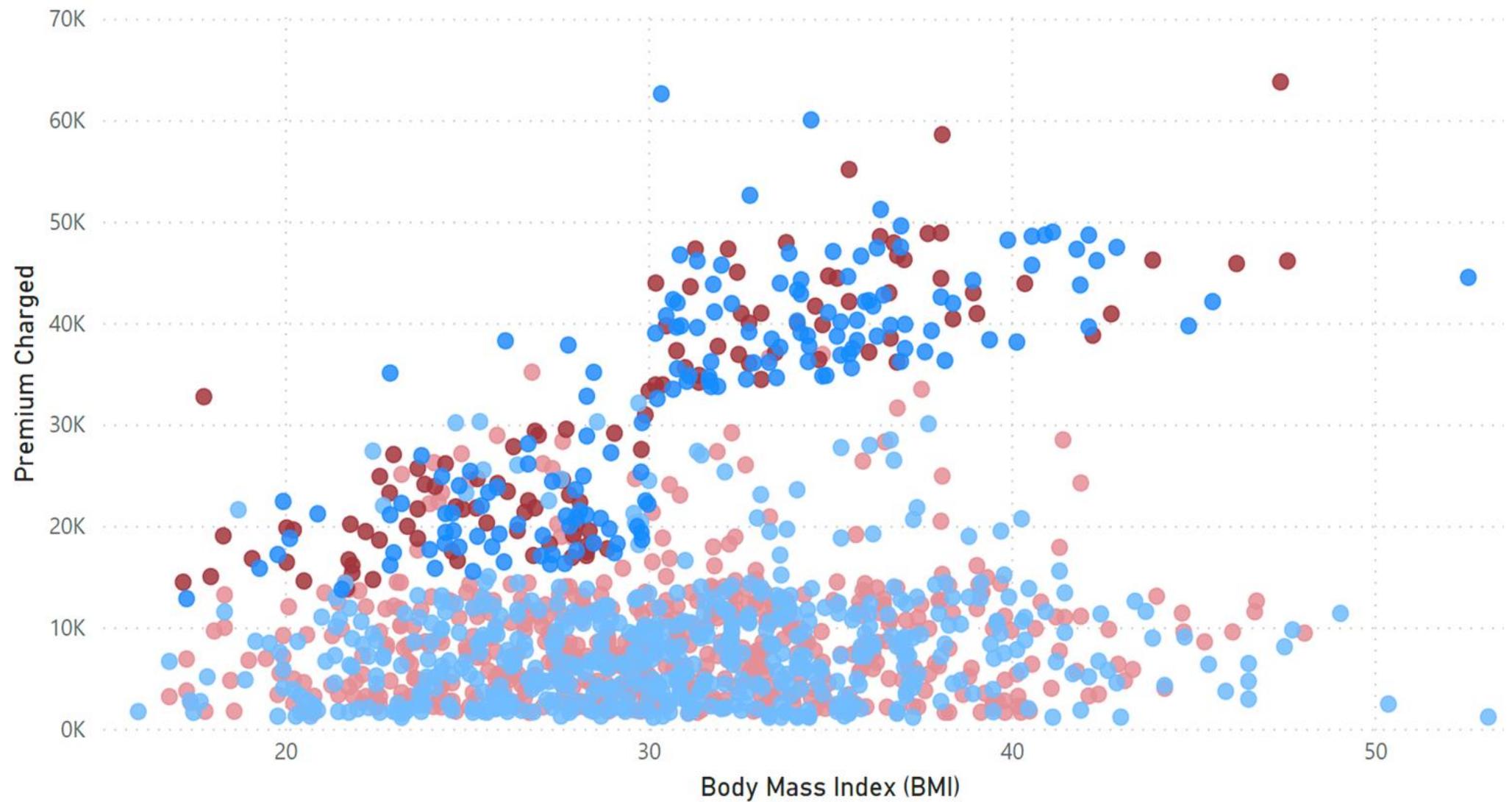


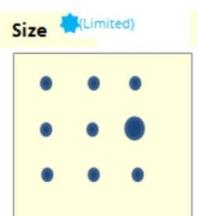
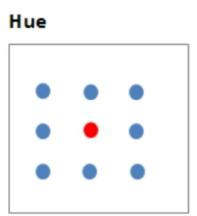
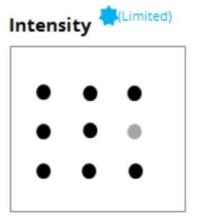
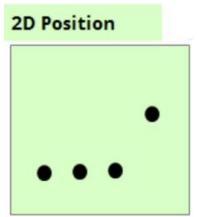
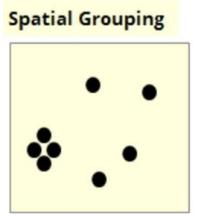
**smoker** ● no ● yes





**Sex & Smoker** ● female-no ● female-yes ● male-no ● male-yes





**Sex & Smoker** ■ female-no ■ female-yes ■ male-no ■ male-yes

70K

60K

50K

40K

30K

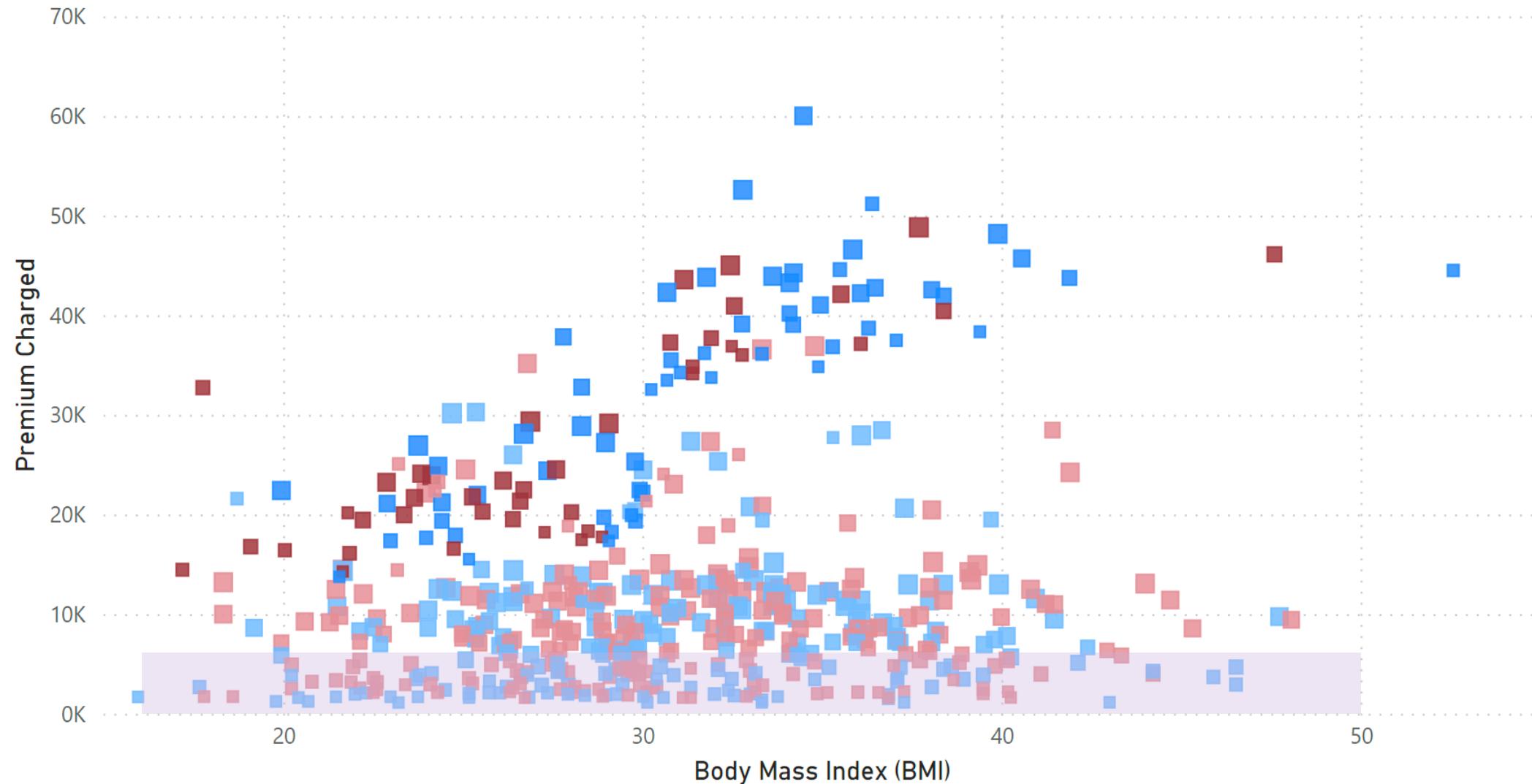
20K

10K

0K

**Body Mass Index (BMI)**

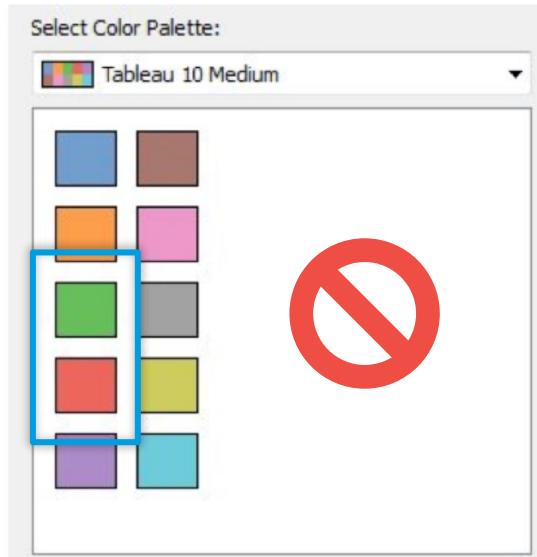
Size indicates Age (18 – 64)



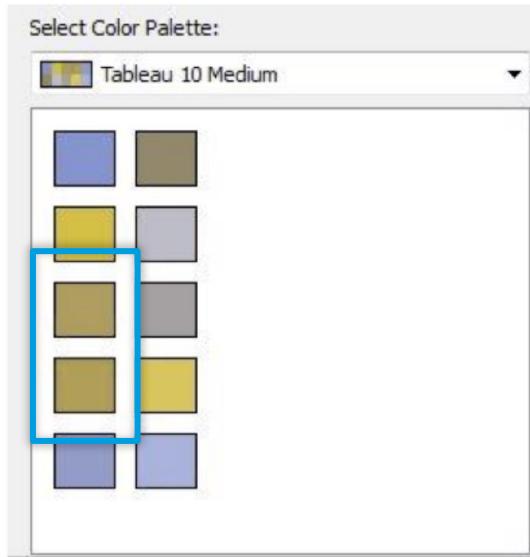
# THINGS TO AVOID

## Red/Green Deutanopia (6% of males)

Original Image

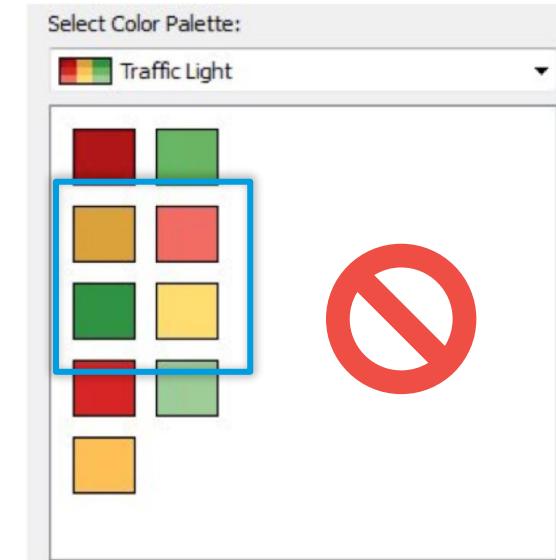


Deutanope Simulation

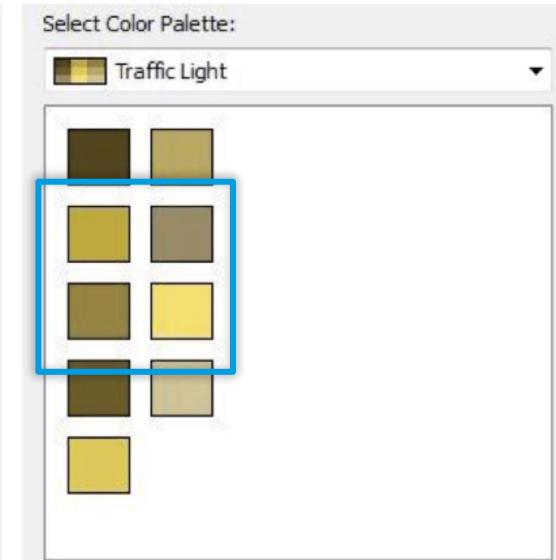


## Red/Green Protanopia (2% of males)

Original Image

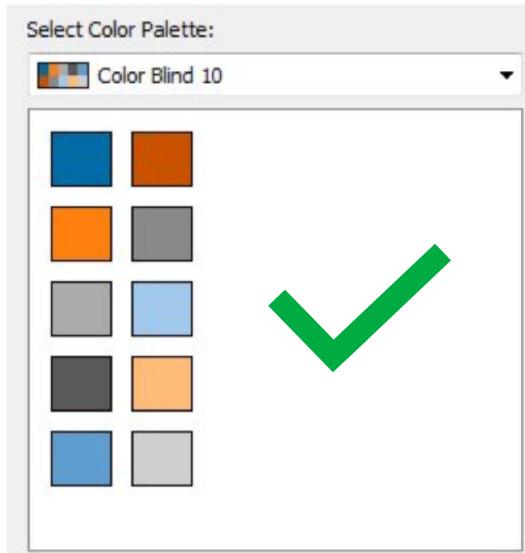


Protanope Simulation

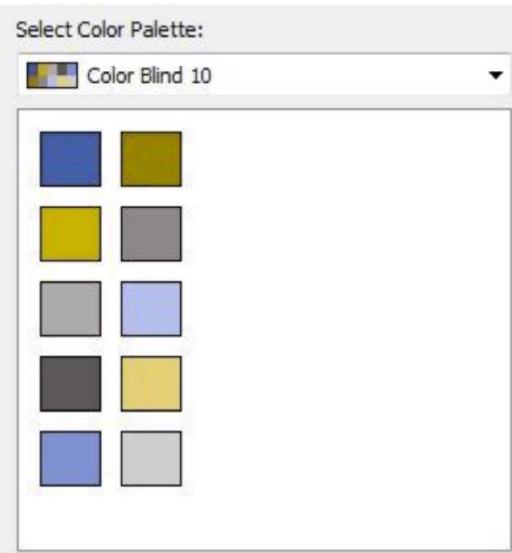


## Red/Green Deutanopia (6% of males)

Original Image

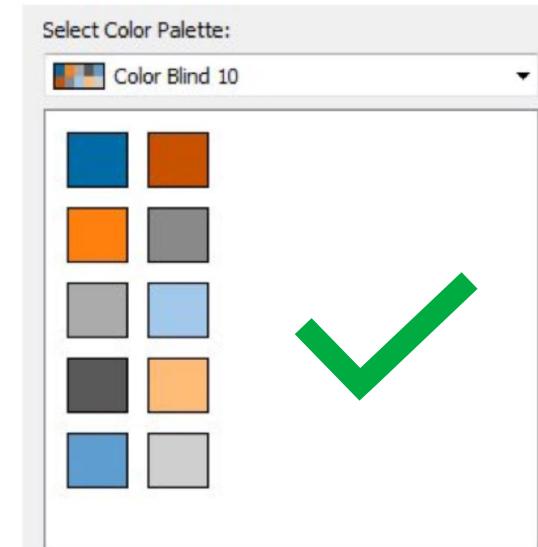


Deutanope Simulation

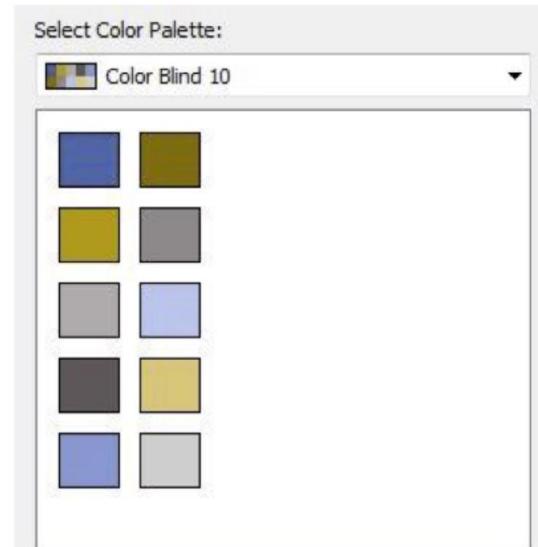


## Red/Green Protanopia (2% of males)

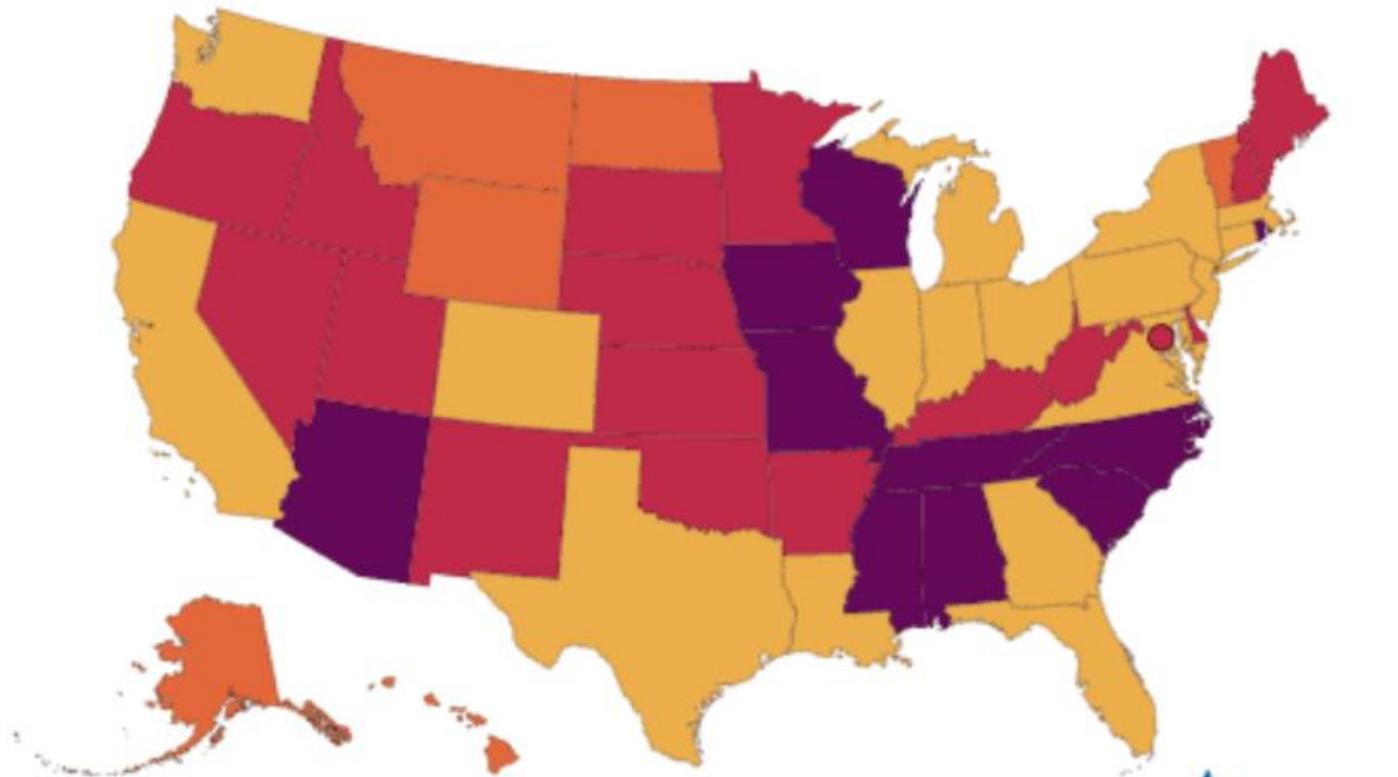
Original Image



Protanope Simulation



18 states report more than 10,000 cases of COVID-19.



AS GU MH FM MP PW PR VI



### Reported Cases

- |                |                 |
|----------------|-----------------|
| 1 to 100       | 10,001 or more  |
| None           | 101 to 1,000    |
| 1,001 to 5,000 | 5,001 to 10,000 |



Graphics that  
are accurate  
but misleading

Baseline should  
start at zero,

... + 77

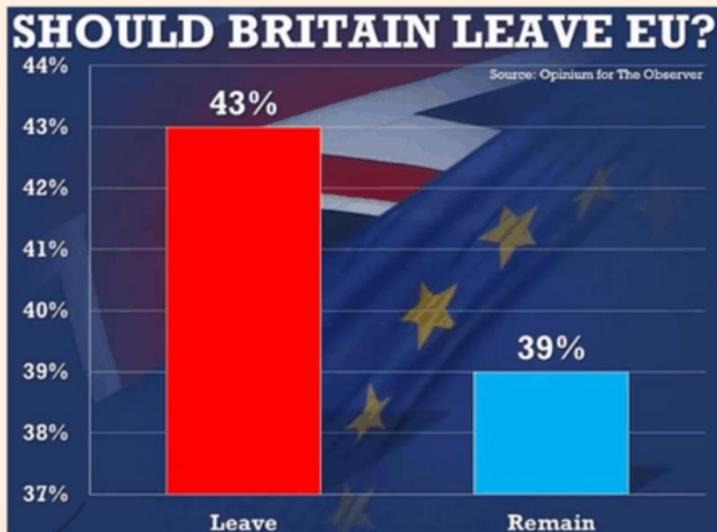
### How Britain voted

Older people with fewer formal qualifications most likely to have voted Leave

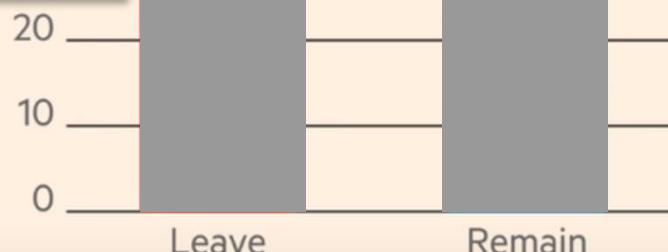
| 2015 vote        |        |       |
|------------------|--------|-------|
|                  | Remain | Leave |
| Conservatives    | 39     | 61    |
| Labour           | 65     | 35    |
| Liberal Democrat | 68     | 32    |
| UKIP             | 5      | 95    |
| Green            | 80     | 20    |

are accurate  
but misleading

A better chart of  
the same data

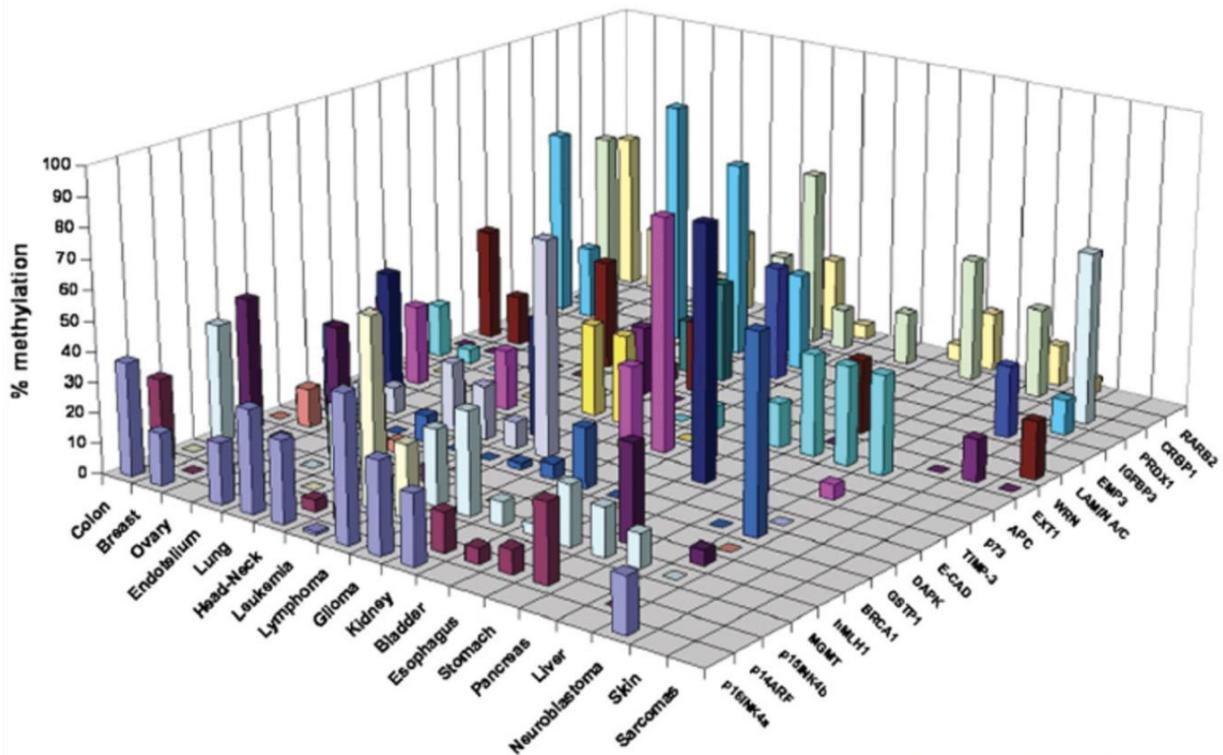


### Should Britain leave the EU?





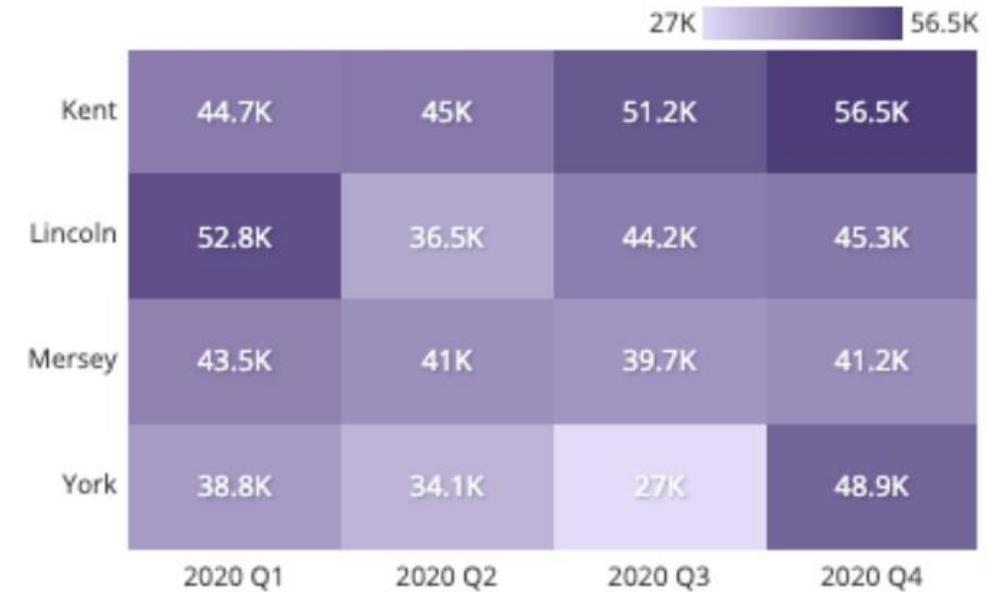
## A CpG Island Hypermethylation Profile of Human Cancer



Hum. Mol. Genet. (2007) 16:R50-59



## New Revenue



# PROBLEMS & TOOLS TO FIX THEM

# I HAVE DATA, BUT I DON'T KNOW WHAT STORY TO TELL.

- Build a predictive model and look at important features.
- Do this quickly with an automated machine learning tool: RapidMiner (point/click), storyteller (code/R), etc.

```
# run model a model to find a story about charges (premiums)
dt %>%
  correlatedfeatures_address(
    target = 'charges'
  ) %>%
  fitmodel() %>%
  summary()

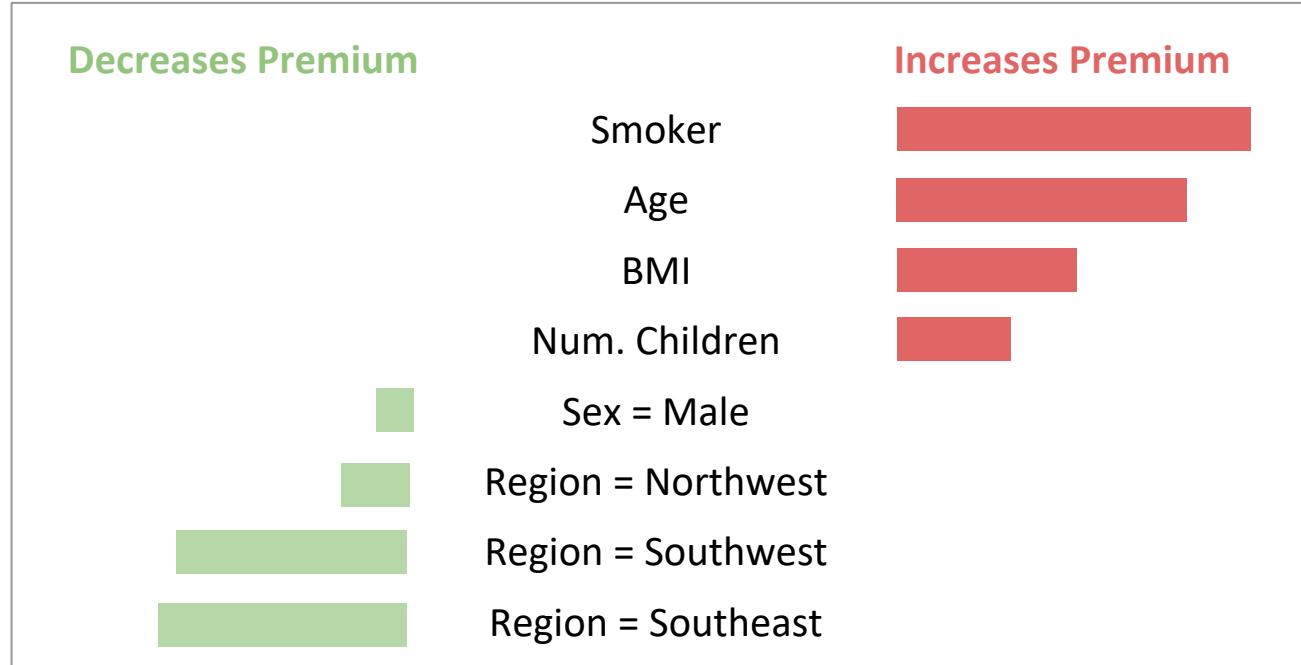
[1] "charges"

Call:
lm(formula = y ~ ., data = yX)

Residuals:
    Min      1Q  Median      3Q     Max 
-10584 -2748 -1068   1092  24373 

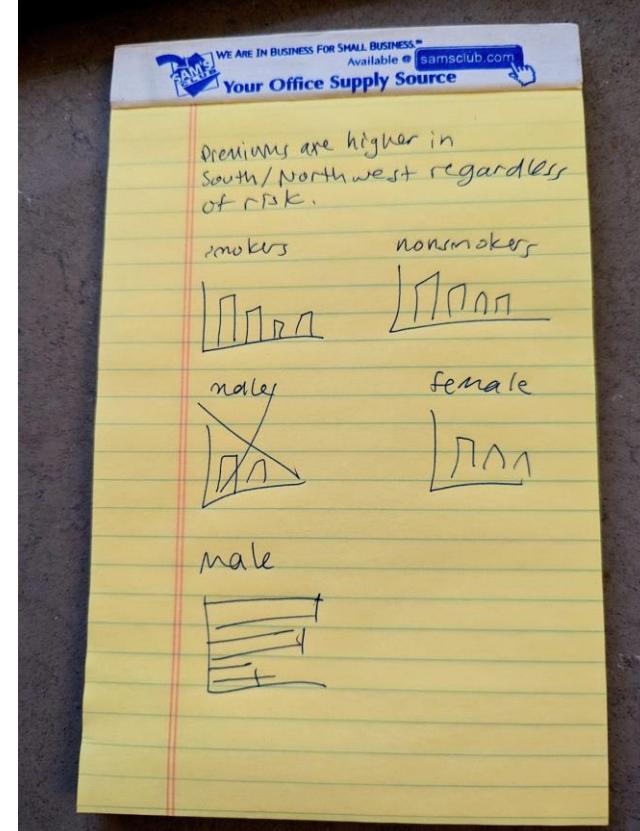
Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept) -11003.70    965.52 -11.397 < 2e-16 ***
age          250.67     11.48  21.841 < 2e-16 ***
bmi          317.37     28.15  11.273 < 2e-16 ***
children     519.40    132.24   3.928 9.03e-05 ***
smokerTRUE   22885.98   403.69   56.692 < 2e-16 ***
sex.male     -106.99    320.02  -0.334  0.7382  
region.southeast -1072.18  460.65  -2.328  0.0201 *  
region.northwest -444.34  456.46  -0.973  0.3305  
region.southwest -1021.20  458.71  -2.226  0.0262 *  
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 5787 on 1309 degrees of freedom
Multiple R-squared:  0.7448,    Adjusted R-squared:  0.7432 
F-statistic: 477.4 on 8 and 1309 DF,  p-value: < 2.2e-16
```



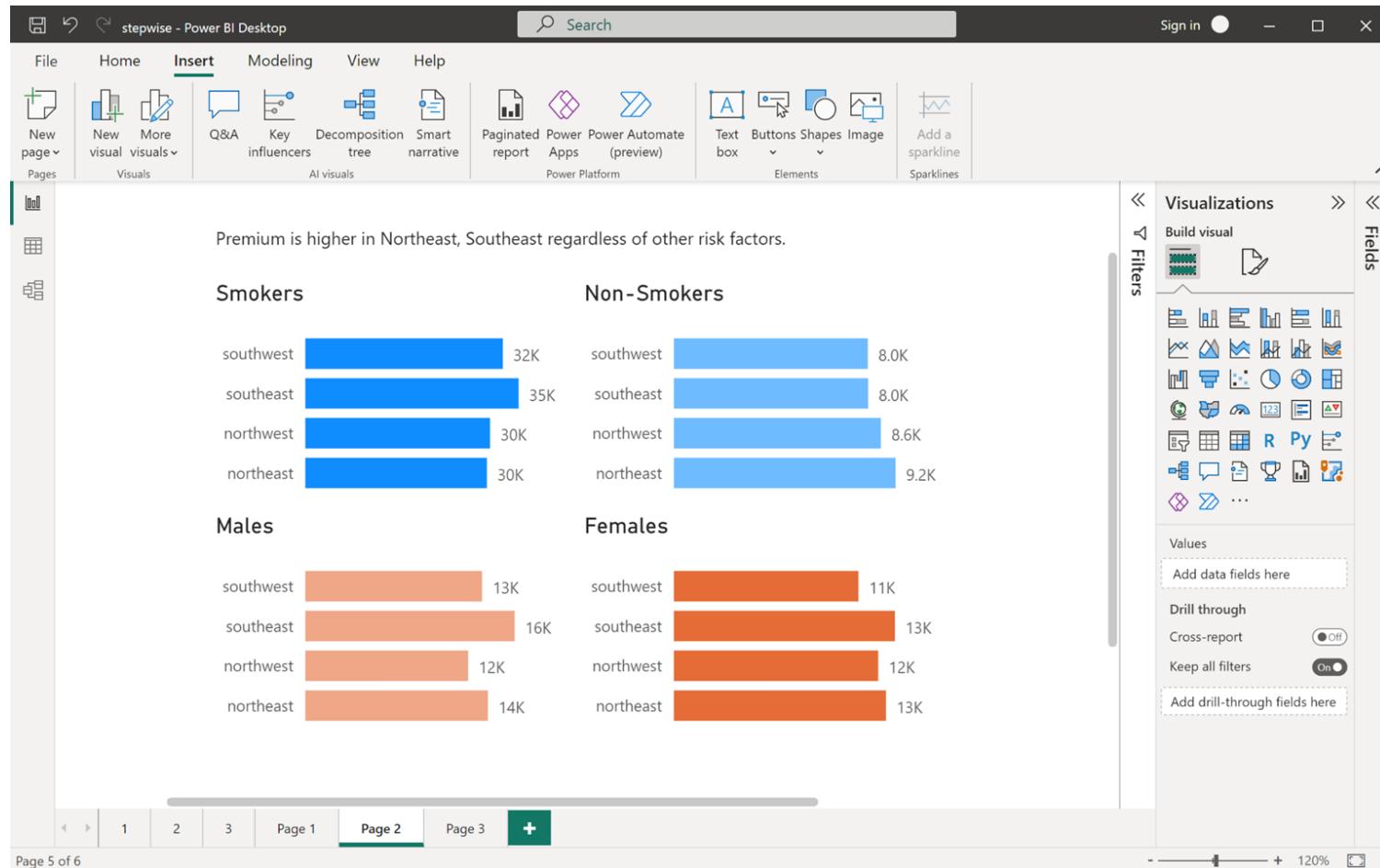
# I KNOW THE STORY I WANT TO TELL, BUT WHAT VIZ SHOULD I USE?

- Get inspiration from the Financial Times Visual Vocabulary
  - Use hand-drawing to quickly prototype an idea.



# I AM READY TO MAKE A VISUALIZATION, BUT I AM SHORT ON TIME.

- Use a click/drag tool: PowerBI Free Desktop, Excel PivotChart, etc.
- Avoid code-based solutions that can take a long time.



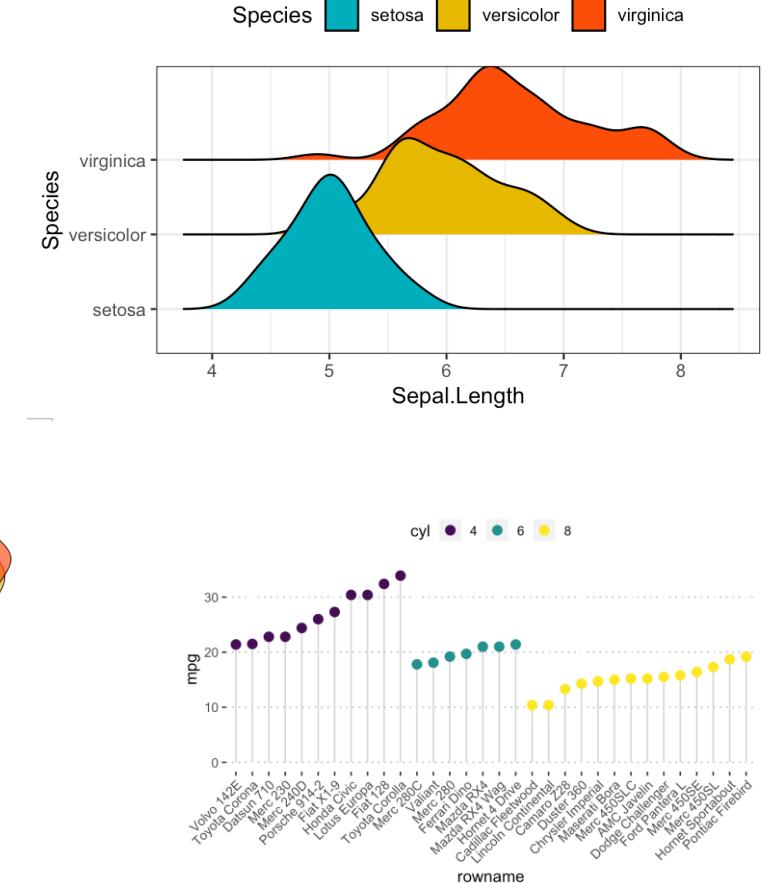
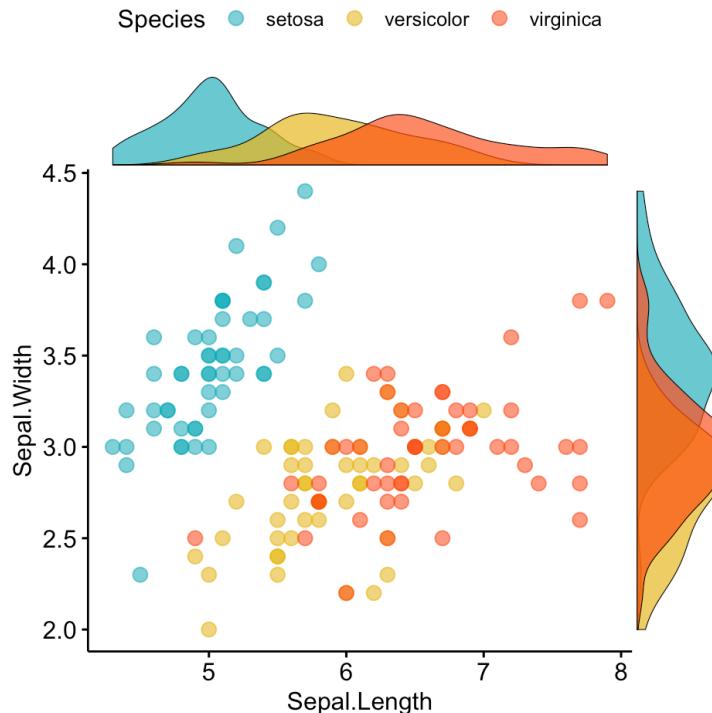
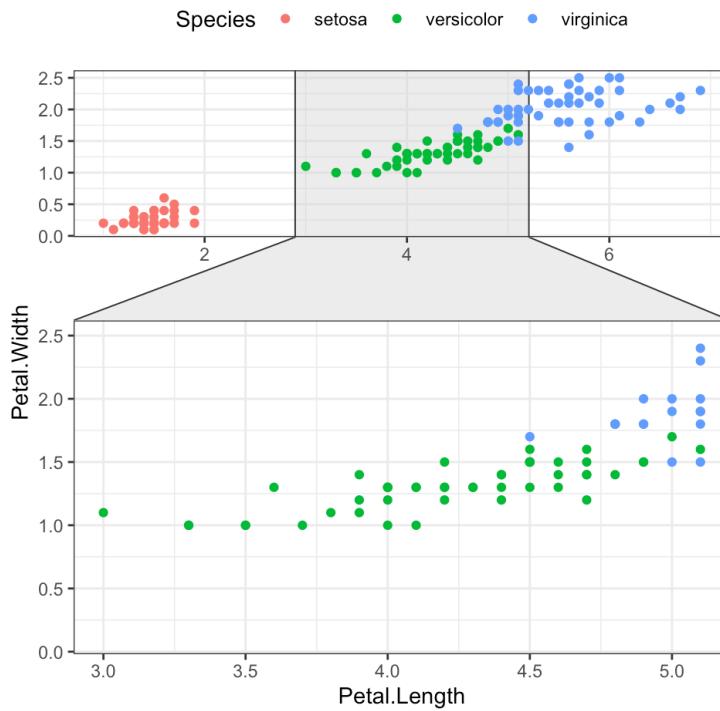
# I'M NOT HAPPY WITH THE DESIGN OPTIONS IN POWER BI

- Export from GUI to PDF and edit with a design tool like Adobe Illustrator, Inkscape, or Fiji.
- This is not intuitive, so watch a video to see how (link on last slide).



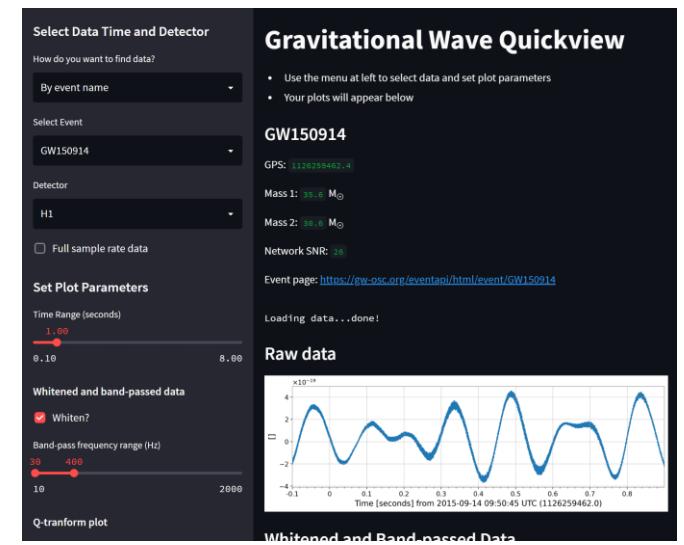
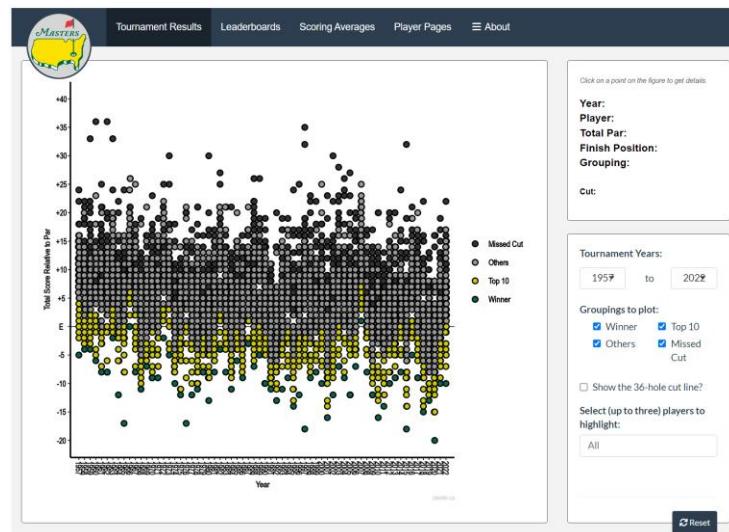
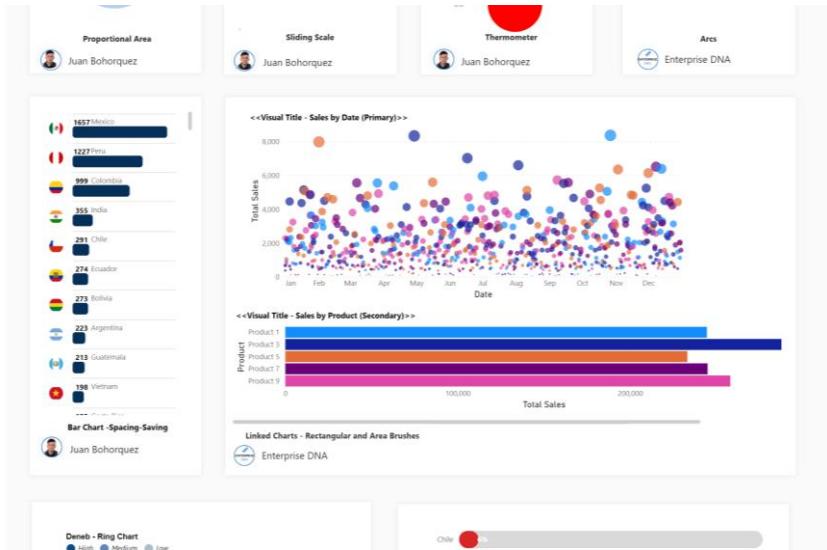
# I AM NOT HAPPY WITH CHART TYPE / NEED TO MAKE MANY SIMILAR CHARTS

- Use a code-based solution so you can automate the process or use new chart types.
- R: ggplot, RMarkdown; Python: seaborn, Jupyter.



# I NEED TO BUILD A DYNAMIC APPLICATION.

- Use a tool you can publish to the web and give users power to filter, etc.
- Power BI, Tableau (click/drag), R Shiny (code, R) or Streamlit (code, Python)



<https://community.powerbi.com/t5/Data-Stories-Gallery/My-own-Gallery/td-p/3054132>

<https://shiny.rstudio.com/gallery/masters.html>

<https://gw-quickview.streamlit.app/>

## AutoML

Let the computer do the work for you.  
Rapidminer, Storyteller

Found your story?

## Free-hand Drawing

Quickly prototype your final product.

Know what you  
need?

## Business Intelligence

Prototype with data.  
PowerBI Desktop, Excel PivotChart

Extra mile,  
one time?

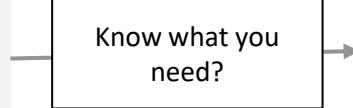
Need to automate?

## Design

Adobe Illustrator

## Code-based

R, R Shiny, Python



# QUESTIONS



**Dalesa Bady, ACAS, MAAA**

Actuary, GuideOne Insurance

[linkedin.com/in/dalesa-bady-acas-05822336](https://linkedin.com/in/dalesa-bady-acas-05822336)



**Bryce Chamberlain, ASA, MSCA**

Principal, Oliver Wyman Actuarial Consulting

[linkedin.com/in/brycechamberlain](https://linkedin.com/in/brycechamberlain)

Get this deck + linked resources at  
<https://bit.ly/gis23-dataviz>



# CAS Student Central Membership Benefits

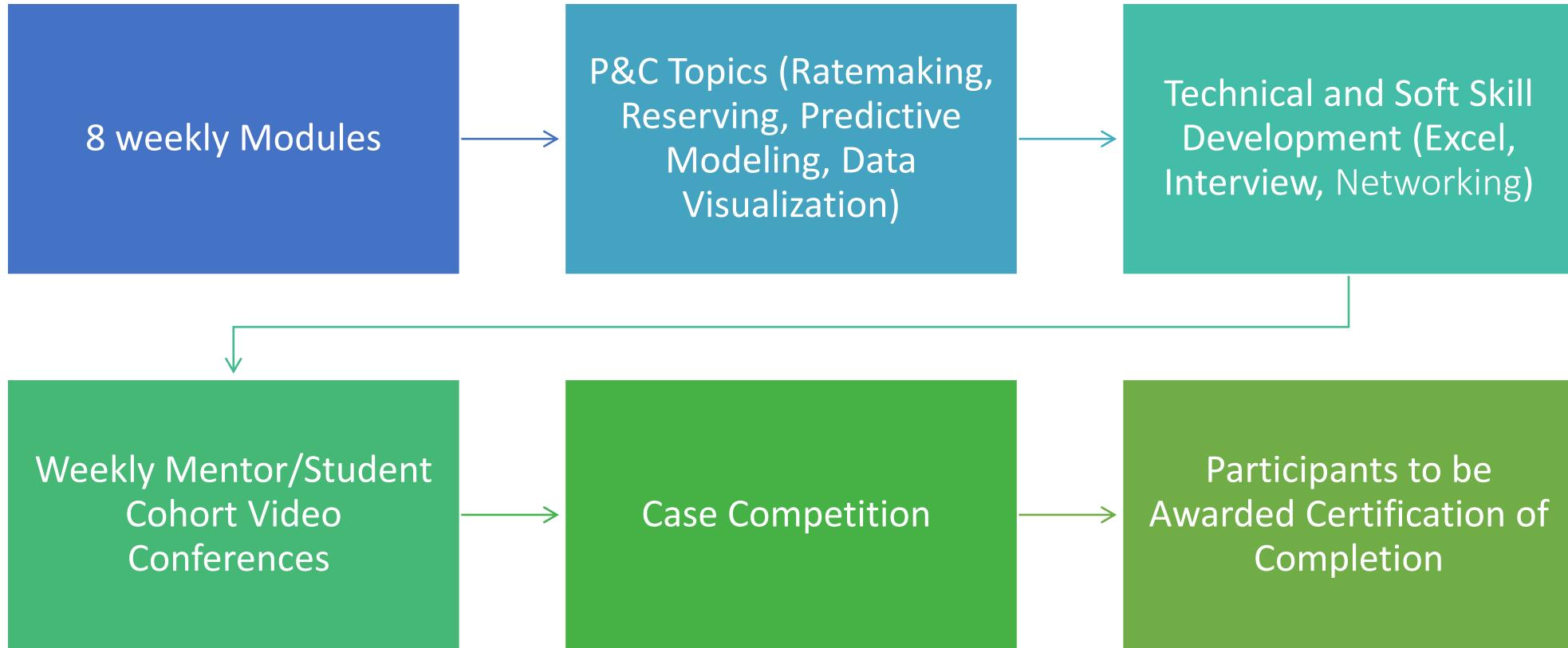


The screenshot shows the homepage of the Student CAS Central website. The header includes links for "STUDENT CAS CENTRAL", "ABOUT THE CAS", "ABOUT OUR PROFESSION", "MEMBER RESOURCES", "NEWS", "Sign in", and a search icon. The main banner features a diverse group of smiling professionals with the text "What is a property and casualty actuary? We are the leading professionals in finding ways to manage risk." Below the banner are two columns: "About the CAS" (describing it as the only actuarial organization focused on property and casualty risks) and "Why Join CAS Student Central" (describing it as a membership program for university students interested in pursuing an actuarial career).



# STUDENT CENTRAL

## SUMMER PROGRAM



## Worksheets

*Calculator*  
Excel, Google Sheets

Most people can use it.  
Easily manipulate single records.

Easy to make mistakes.  
Difficult to automate.  
Slow on large data.

## Business Intelligence

*Get started now*  
Power BI, Tableau, PivotCharts

Lots of options quickly.  
Click & drag

Limited functionality.  
Difficult to automate.

## AutoML

*Search for insights*  
RapidMiner, Storyteller

Find stories across all data.

Limited visualization.  
Results are complex.  
Expensive if not open source.

## Design

*Make it pretty*  
Adobe Illustrator, Inkscape

Lots of features and options for perfecting the visual.

Very time consuming.  
Software is complex, difficult to learn.

## Free-hand Drawing

*Begin with the end in mind*  
Pen & Paper, Tablet

Fastest method, no interface to slow you down.

Not generated by data.  
Not fit for delivery.

## Code-based

*Automate repetitive tasks*  
RMarkdown, R Shiny

Unlimited functionality.  
Open source.  
Git version control.

Time consuming.  
Need to code.