Original Proposed Content

Title: Effective Data Visualization for Actuaries

Session Description:

Actuaries often need to regularly communicate trends, diagnostics, and findings to stakeholders (other actuaries, their business partners, or clients).

Two questions arise:

- How do we create visualizations that let actuaries focus on interpreting rather than doing?
- How do we re-structure these visualizations to tell a better story?

In this session Brian and I will examine some common actuarial analysis problems (e.g. loss development, trend/inflation, mix shift exhibits, and model diagnostics) through the lens of visualization. Using R's `ggplot2` package, we'll show how these visualizations can be improved to better communicate the story. For example, reserve loss development can be manipulated with `dplyr` to show cumulative losses on the x-axis and incremental amounts on the y-axis to better highlight the loss development factor. After demoing a few visualization examples, we'll show how some formatting tricks can be used to make your visualizations look polished.

Learning Objectives:

1. Demonstrate when alternative representations of common actuarial visualizations are helpful
2. Basic use of the `ggplot2` package
3. When and how to use non-data formatting to customize a visualization for your audience

Session Format: Concurrent Session, 50-60 minutes

Theme: Navigate

Applicable Areas: Actuarial Toolkit: Programming Software, Data Manipulation Tools, and Data Visualization Tools

Knowledge Level: Level 2: General Knowledge of the Subject (6-9 years)

Interactive Elements: There will be one or two discussion prompts, as well live voting to compare different means of visualizing data analysis

To Be Removed. For Reference Only.

Ideas for Interactivity

- **Familiarity with R**
- **“Which chart is easier to interpret”**
- **Test Bar/Line bias (**https://bit.ly/3ymQ1eE**)**

Effective Data Visualization for Actuaries

Brian Fannin

Jordan Bonner

Antitrust Notice

- **The Casualty Actuarial Society is committed to adhering strictly to the letter and spirit of the antitrust laws. Seminars conducted under the auspices of the CAS are designed solely to provide a forum for the expression of various points of view on topics described in the programs or agendas for such meetings.**
- **Under no circumstances shall CAS seminars be used as a means for competing companies or firms to reach any understanding – expressed or implied – that restricts competition or in any way impairs the ability of members to exercise independent business judgment regarding matters affecting competition.**
- **It is the responsibility of all seminar participants to be aware of antitrust regulations, to prevent any written or verbal discussions that appear to violate these laws, and to adhere in every respect to the CAS antitrust compliance policy.**

Agenda

- **Data Storytelling**
- **Actuarial Visualization**
- **Automating Actuarial Exhibits**

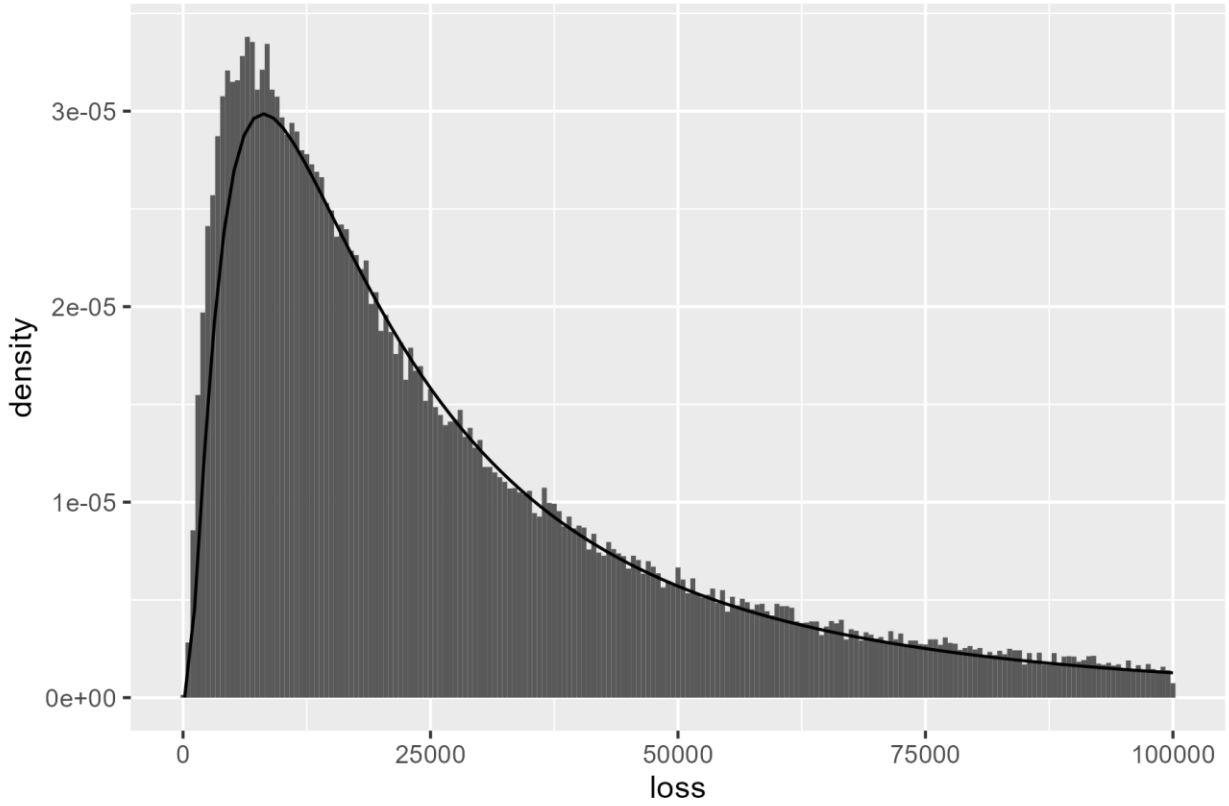
DATA STORYTELLING

Visualizations should:

- **Be Easily Interpreted**
- **Tell a Story**

Initial Example: Fitted Severity

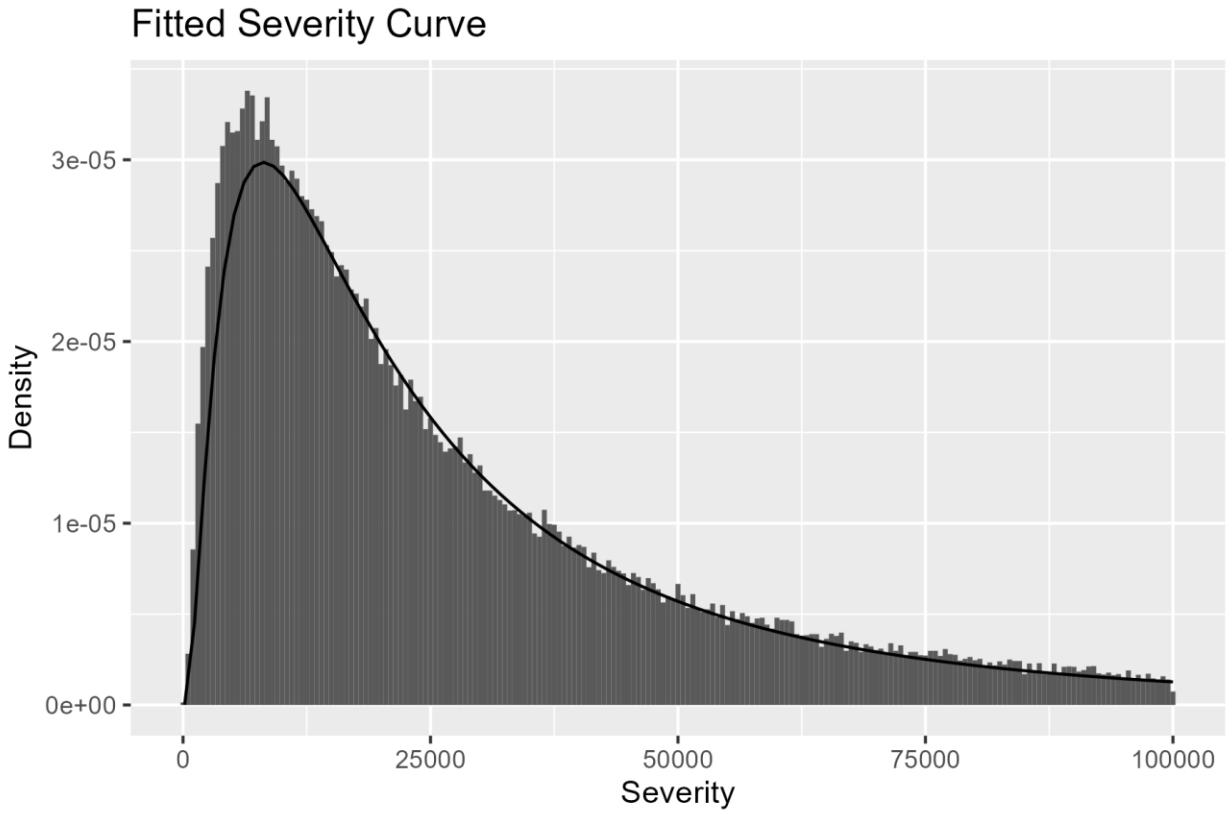
How can we improve upon this chart?




```
p <- data |>
  geom_histogram(aes(x = loss, after_stat(density)) +
  stat_function(fun = dlnorm,
                args = list(meanlog = 10, sdlog = 1.0))
```

Titles & Renaming Variables

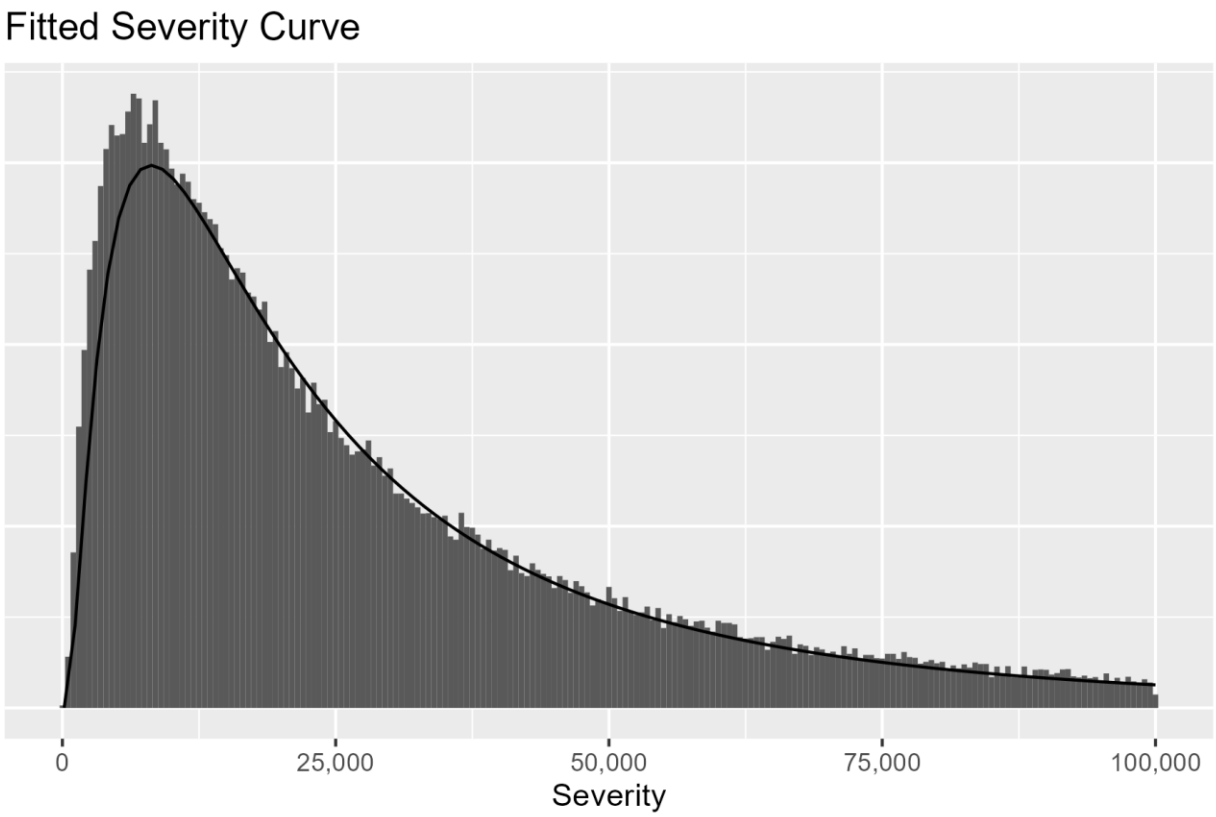
Adding a title and variable names goes a long way.




```
p <- p +
 labs(
    x = "Severity",
    y = "Density",
    title = "Fitted Severity Curve"
  )
```

Cleaning Axes

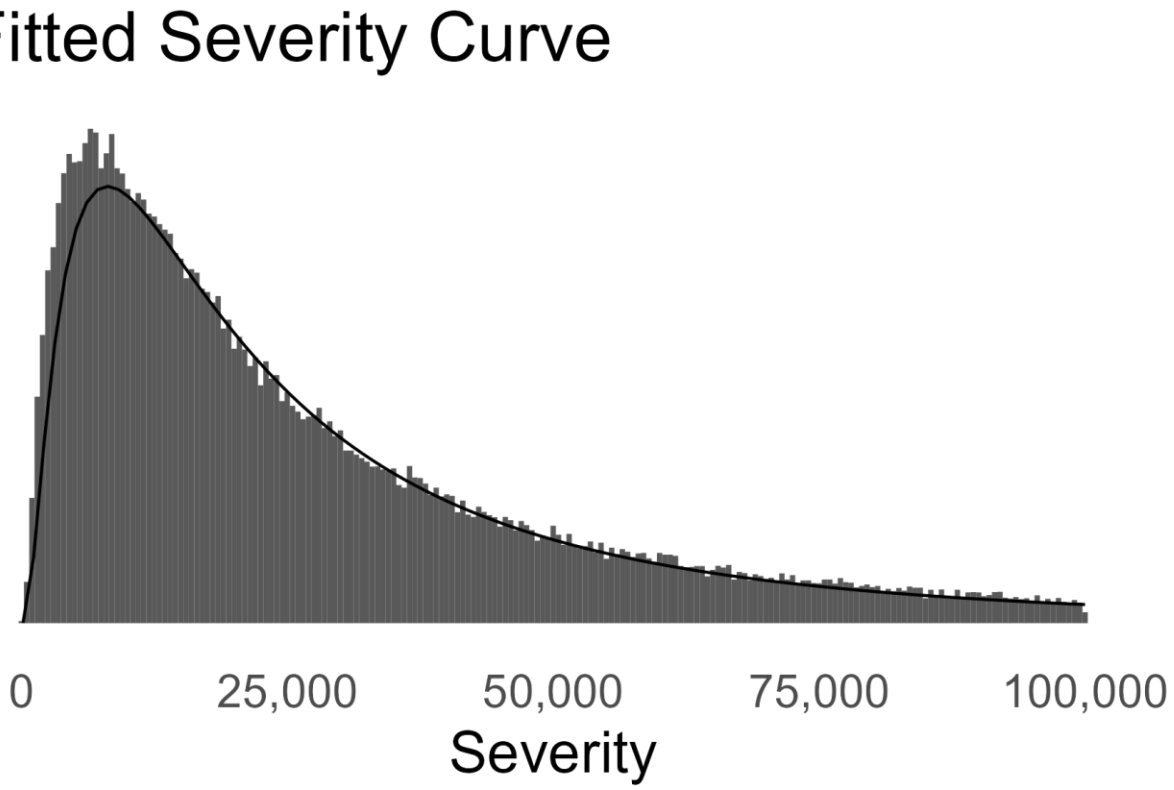
We can change the x-axis labels and remove the y-axis.




```
p <- p +
  scale_x_continuous(labels = scales::comma) +
  theme(axis.text.y = element_blank(),
        axis.ticks.y = element_blank(),
        axis.title.y = element_blank())
```

Text Size & Theme

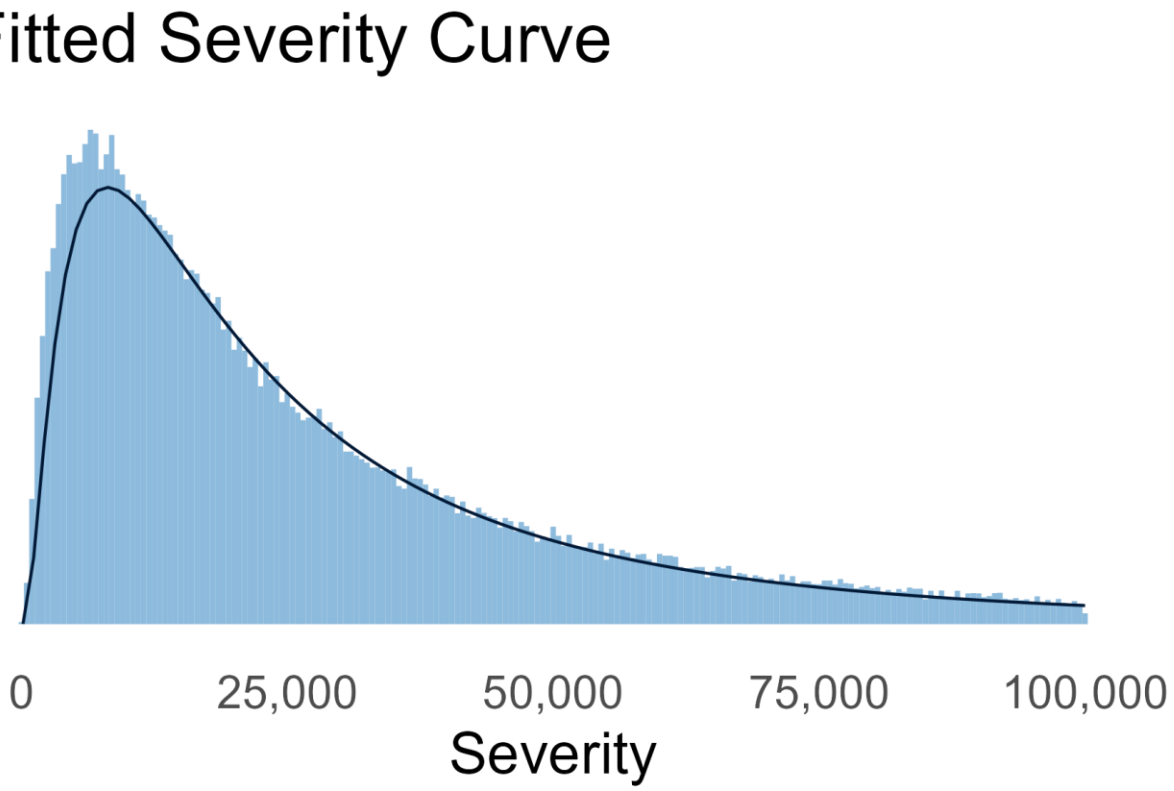
We can increase text size and select a simpler theme.




```
p <- p +
  theme_minimal(base_size = 20) +
  theme(panel.grid.major = element_blank(),
        panel.grid.minor = element_blank()) +
  scale_x_continuous(labels = scales::comma) +
  theme(axis.text.y = element_blank(),
        axis.ticks.y = element_blank(),
        axis.title.y = element_blank())
```

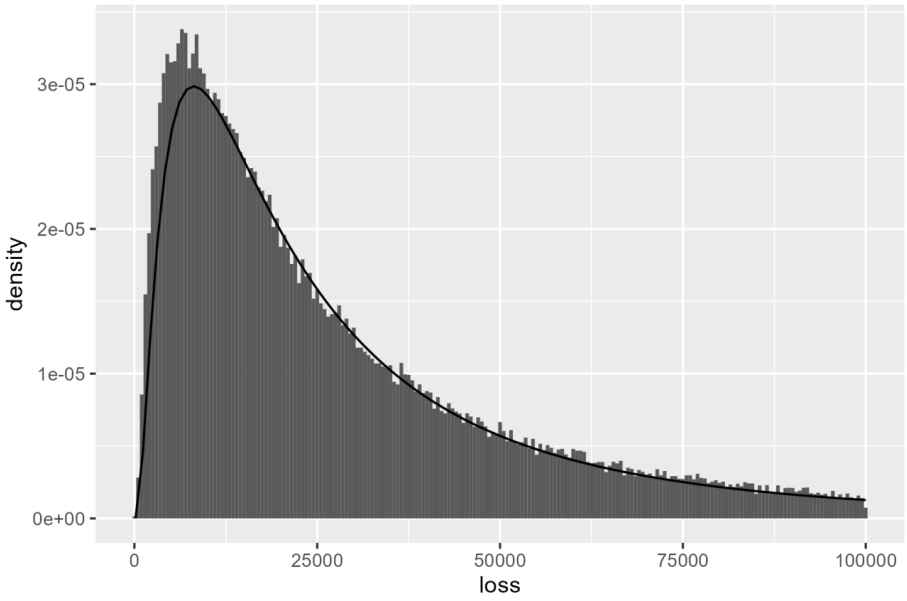
Color

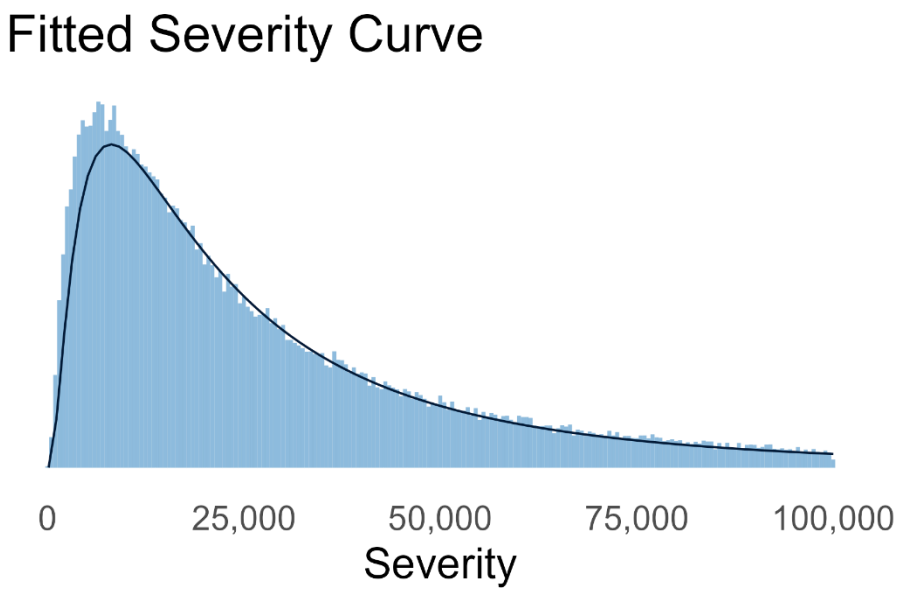
We can also update the color, if desired.



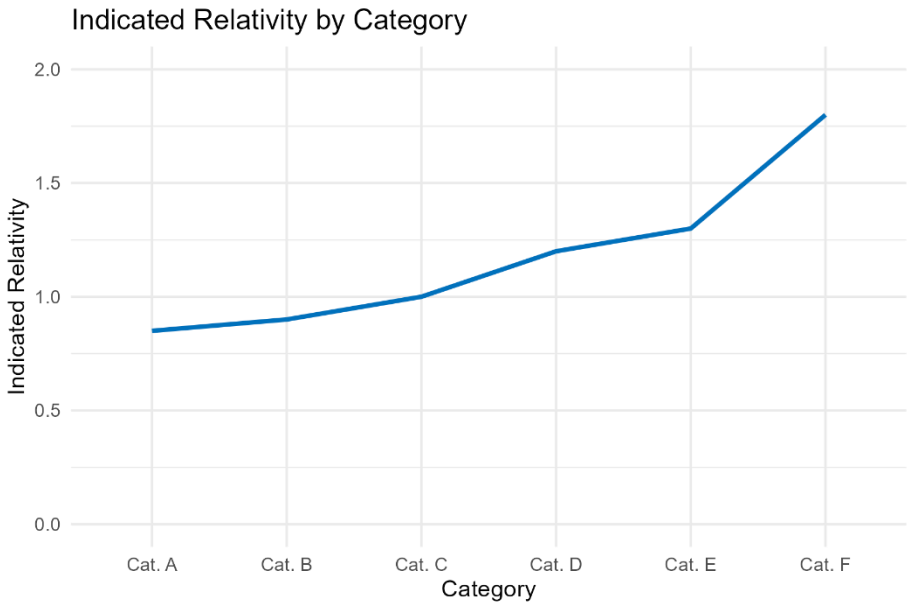

```
...
  geom_histogram(aes(x = loss, after_stat(density)),
                 fill = pal_CAS["light_blue"]) +
  stat_function(fun = dlnorm,
                args = list(meanlog = 10, sdlog = 1.0),
                color = pal_CAS["dark_blue"]) +
...
```

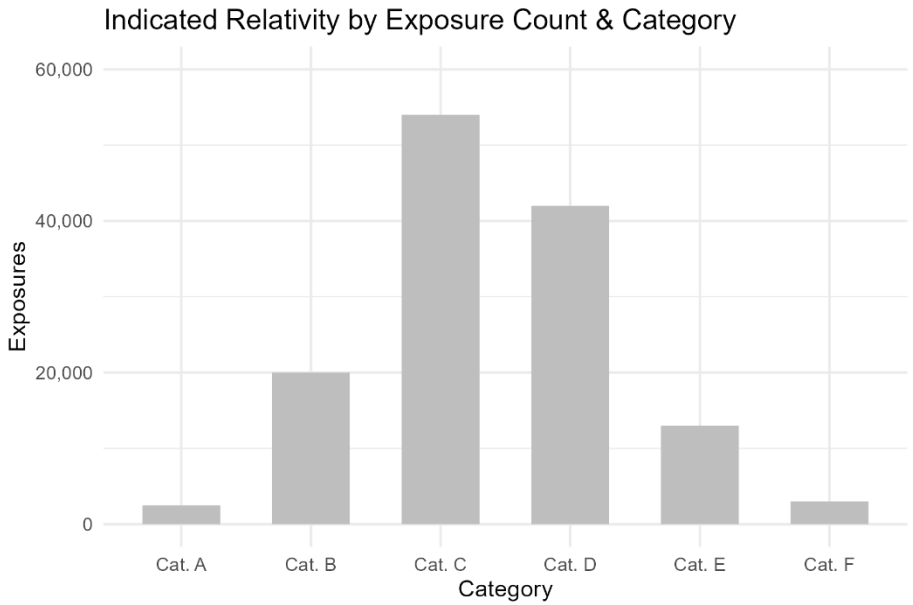
Comparison


3e-05
2e-05
1e-05
0e+00
density
0
25000
50000
75000
100000
loss

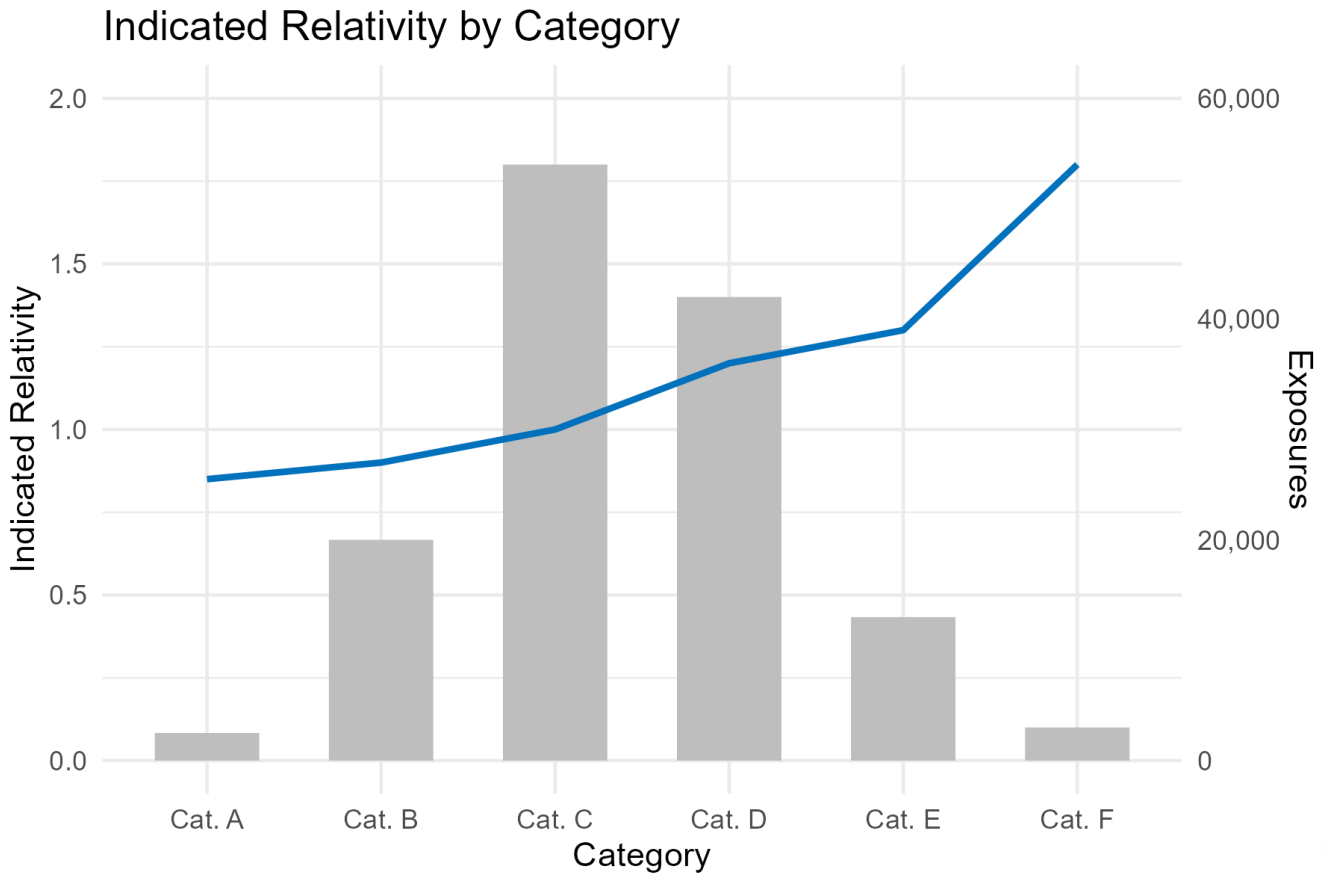

Fitted Severity Curve
0
25,000
50,000
75,000
100,000
Severity

Two Axes vs Two Charts





Two Axes vs Two Charts



ACTUARIAL VISUALIZATION

Common Actuarial Exhibits

- Loss Development Triangles
- Trend Selection
- Rate Indications
- Mix Shift Exhibits
- Model Diagnostics

Loss Development Triangles

Loss Development Triangles

Trend Selection

Trend Selection

Rate Indications

Rate Indications

Mix Shift

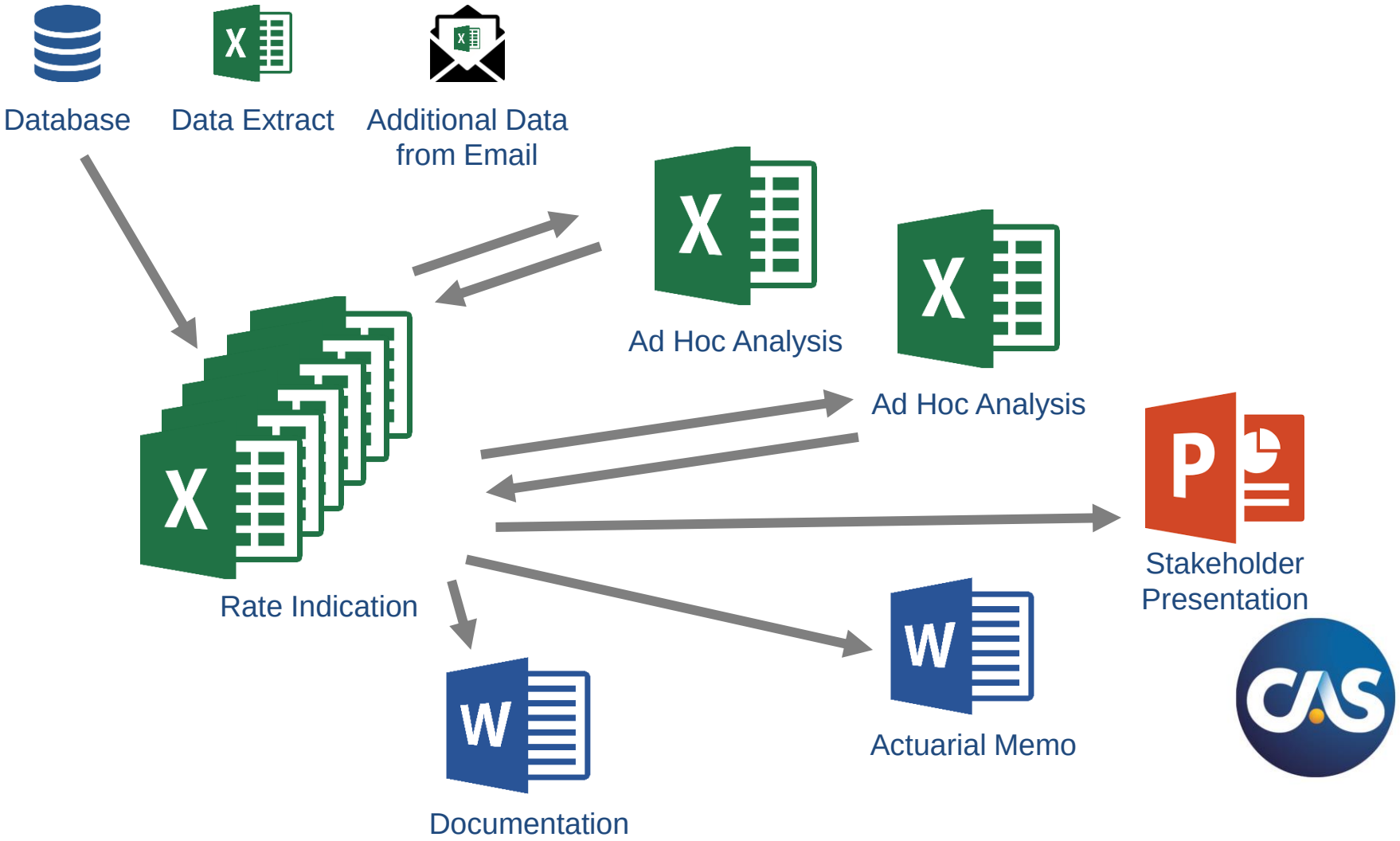
Mix Shift

Model Diagnostics

Model Diagnostics

AUTOMATING ACTUARIAL EXHIBITS

Typical Actuarial Analysis Pipeline

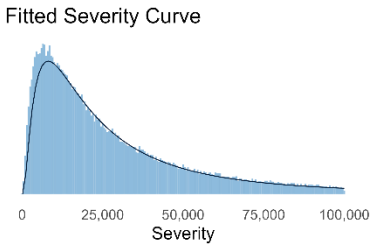


Simplified Actuarial Analysis Pipeline

With rmarkdown:

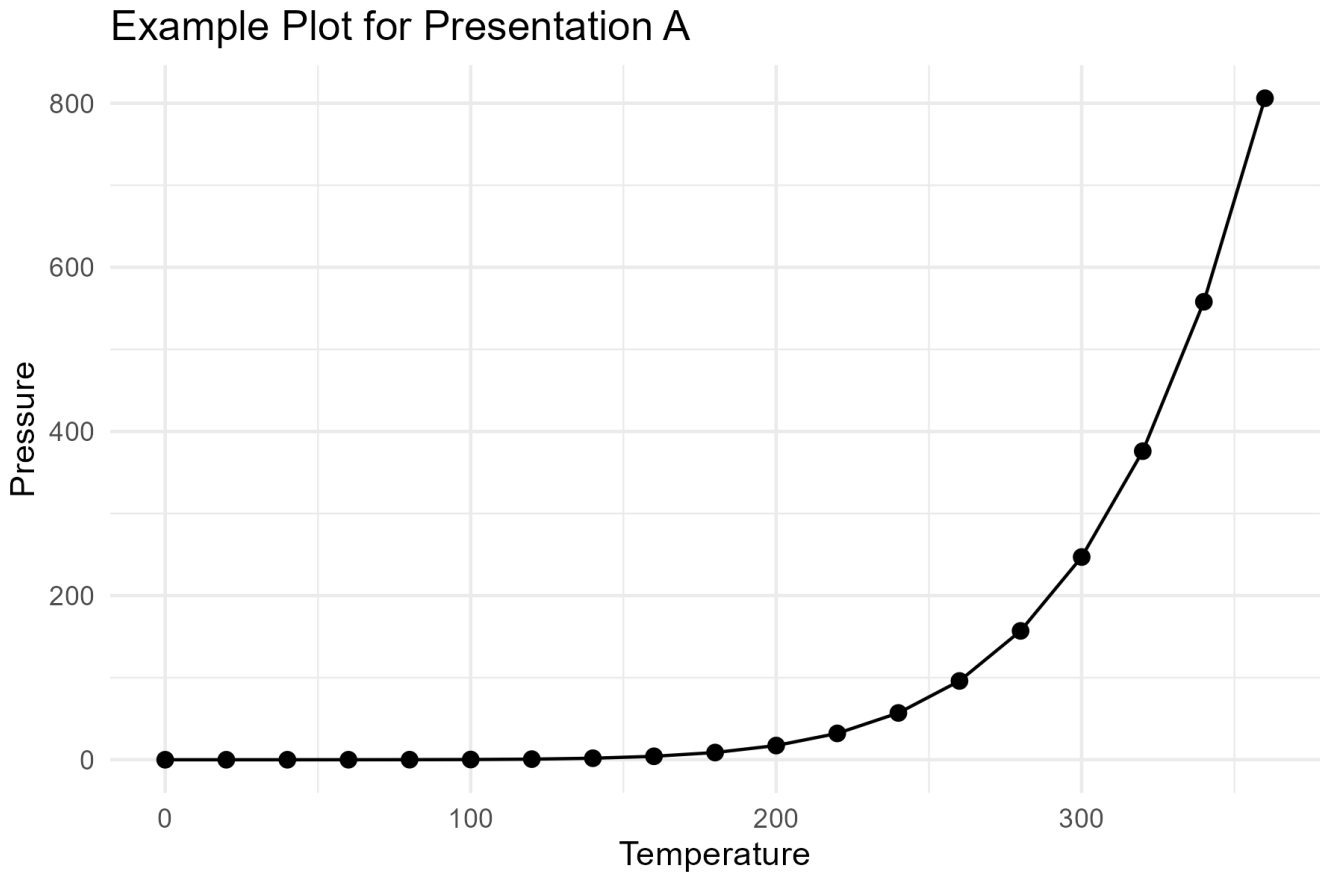
Read in data directly from source
Separate “chunks” for each analysis
Self-documents
New data can flow through
Can also be used to create dynamic slides for presentations
All analysis kept in one place

Database

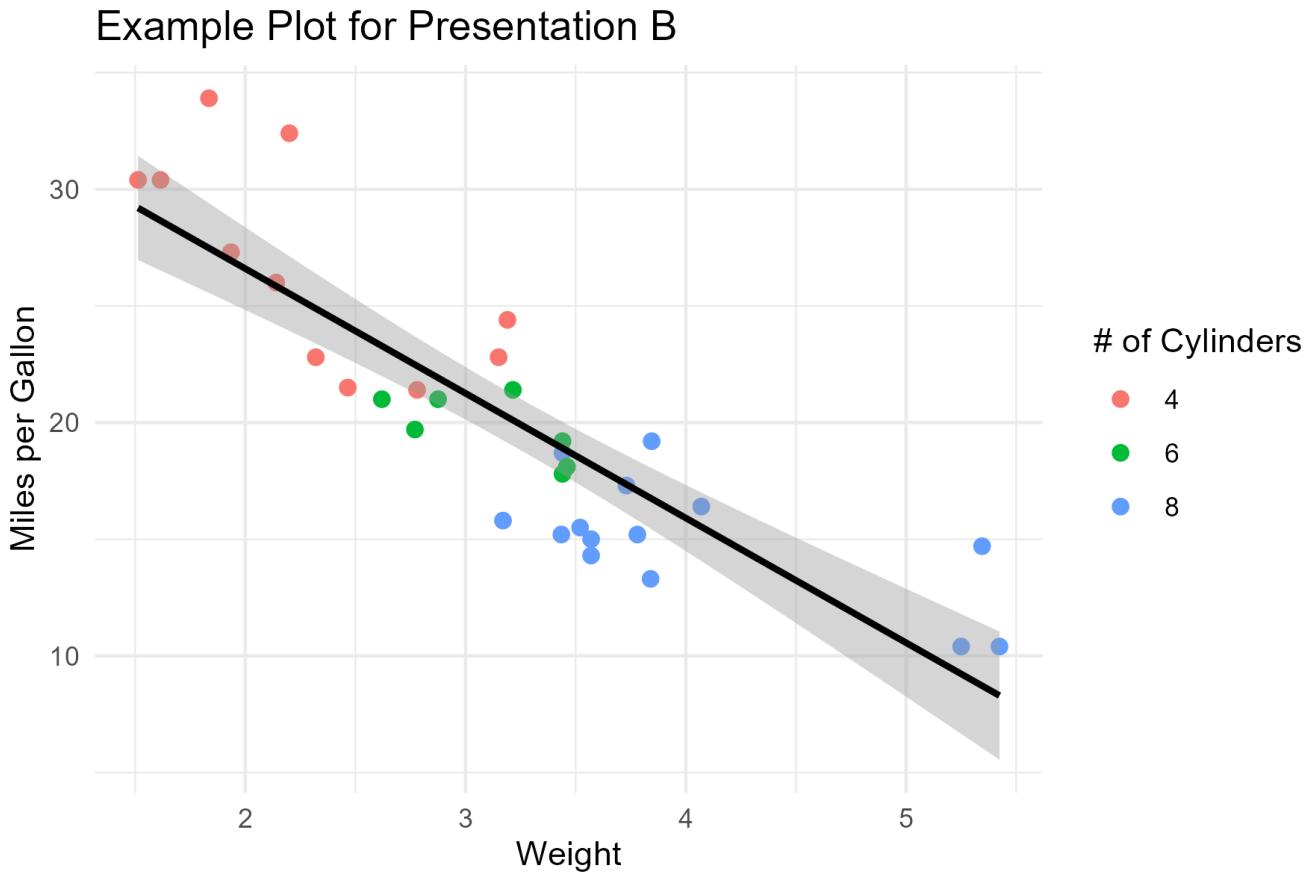
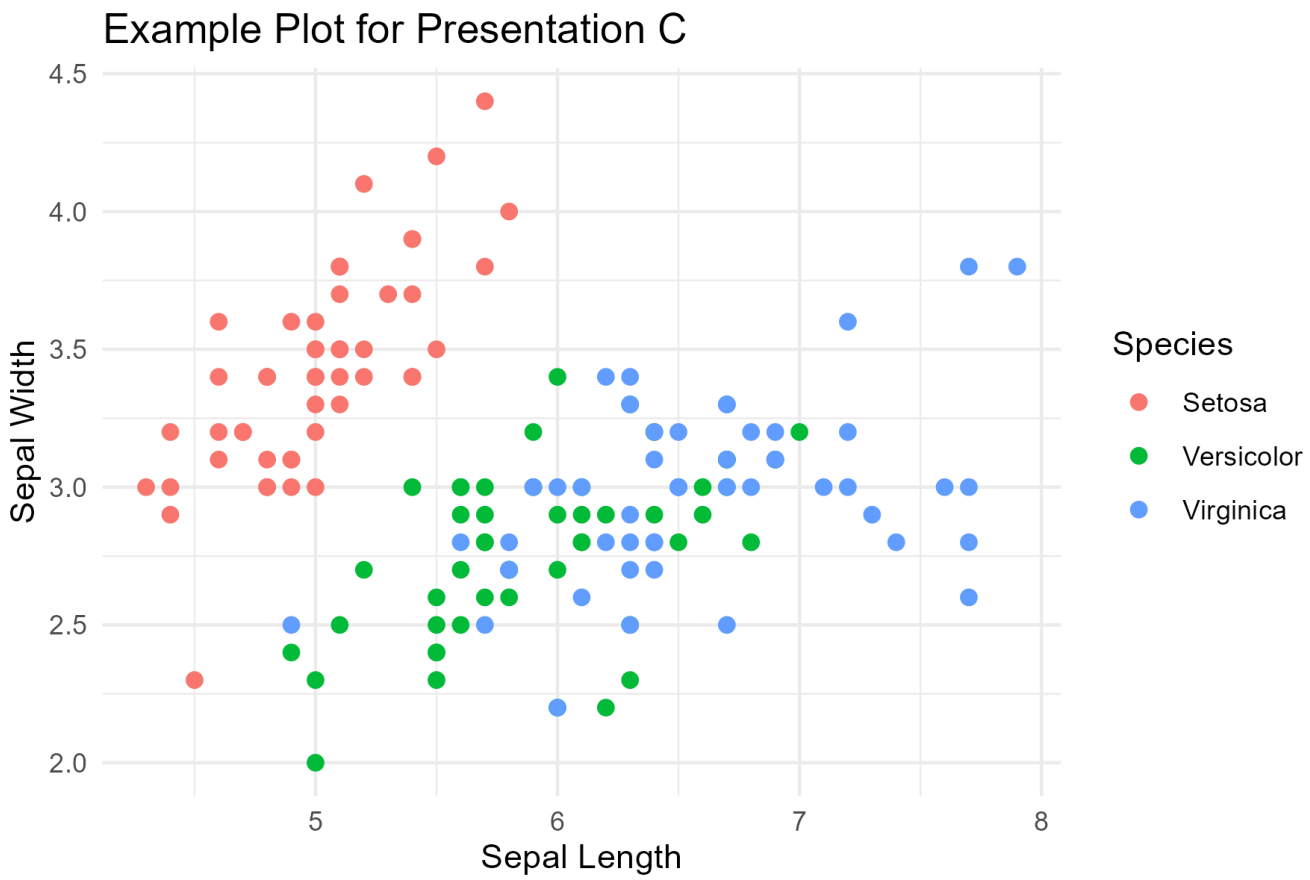


APPENDIX

Example Plot A



Example Plot B



Example Plot C