2018 Predictive Analytics Symposium

Session 35: AP - Practical Aspects of Predictive Models

SOA Antitrust Compliance Guidelines
SOA Presentation Disclaimer

2019 PREDICTIVE ANALYTICS SYMPOSIUM

Session 35: Practical Aspects of Predictive Models

Talex Diede
9/19/19

SOCIETY OF ACTUARIES®

Setting yourself up for success

SOCIETY OF ACTUARIES®


VICTOR QUIT. I NEED YOU TO TAKE OVER HIS PROJECT.
DID HE LEAVE ANY DOCUMEN— TATION?
NO, BUT IT'S OBVIOUSLY SOME SORT OF GLOWING BOX. THAT SHOULD BE ENOUGH TO GET YOU STARTED.
CAN YOU FINISH IT BY MONDAY?
IF IT'S A NIGHT— LIGHT, I CAN FINISH IT BY TODAY.
Dilbert.com DilbertCartoonist@gmail.com
2-27-10 ©2010 Scott Adams, Inc./Dist. by UFS, Inc.


VICTOR DIDN'T LEAVE US MUCH DOCUMENTATION ON HIS PROJECT.
FZEEET!
I GUESS THAT'S WHAT HE MEANT BY "STILL WORKING ON THE GOAT HEAD ISSUE."

Topics

- Reproducible research
  - Documentation
  - Version control
  - Technology
- Peer review

Reproducible Research

The term *reproducible research* refers to the idea that the ultimate product of academic research is the paper along with the laboratory notebooks and full computational environment used to produce the results in the paper such as the code, data, etc. that can be used to reproduce the results and create new work based on the research

- wikipedia

Documentation

- Project Plan and Quality Control for Predictive Analytics Projects.docx
- PM_Review_Checklist.xlsx

Documentation – Project plan

- At inception of project
- After receiving data
- Business problem definition (Scope of the analysis)
- Document data preparation
- Document quality control
- Model form identification
- Establish modeling plan
- Modeling plan execution*
- Quality control
- Deliverables
- Project wrap up

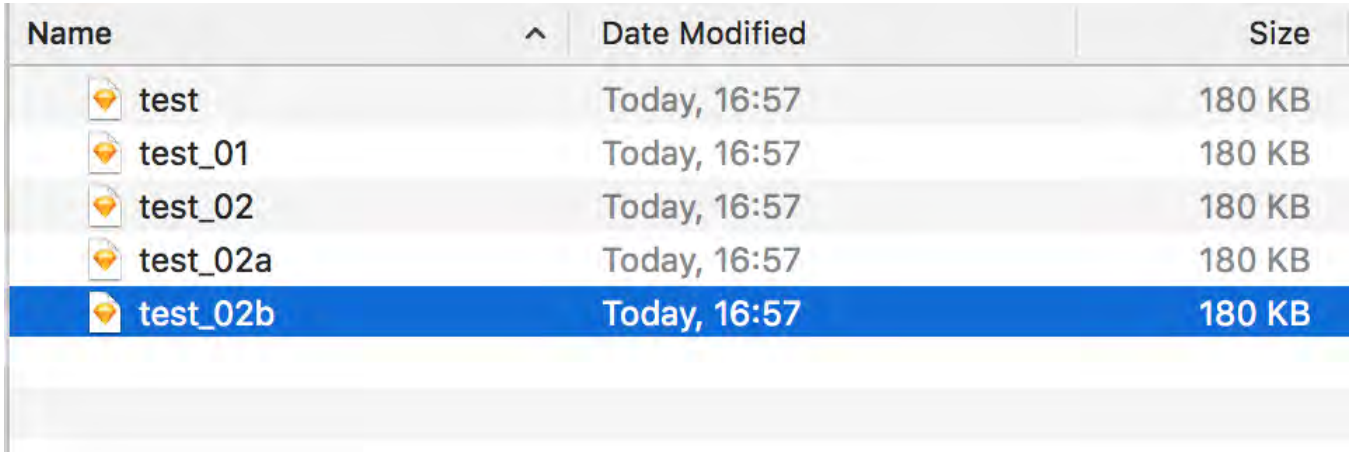
Modeling plan execution

- Create modeling dataset
- Perform variable preselection
- Variable selection
- Variable fit
- Finalize model
- Validate model
- Document final model(s) and datasets
  - Add location and name of script(s), datasets, and models(s) to PMReviewChecklist.xlsx

Version control

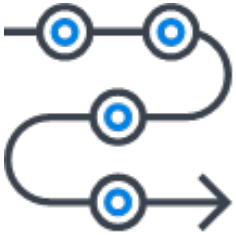
- Category of software tools that help a software team manage changes to source code over time
- Keeps track of every modification to the code in a special kind of database
- If a mistake is made, developers can turn back the clock and compare earlier versions of the code

Version control



Version control_v2

- Utilize a version control system
- Don’t reinvent the wheel
- Make your job easier
- Don’t lose code
- Make it easy to find specific versions
- Be able to roll back to stable state

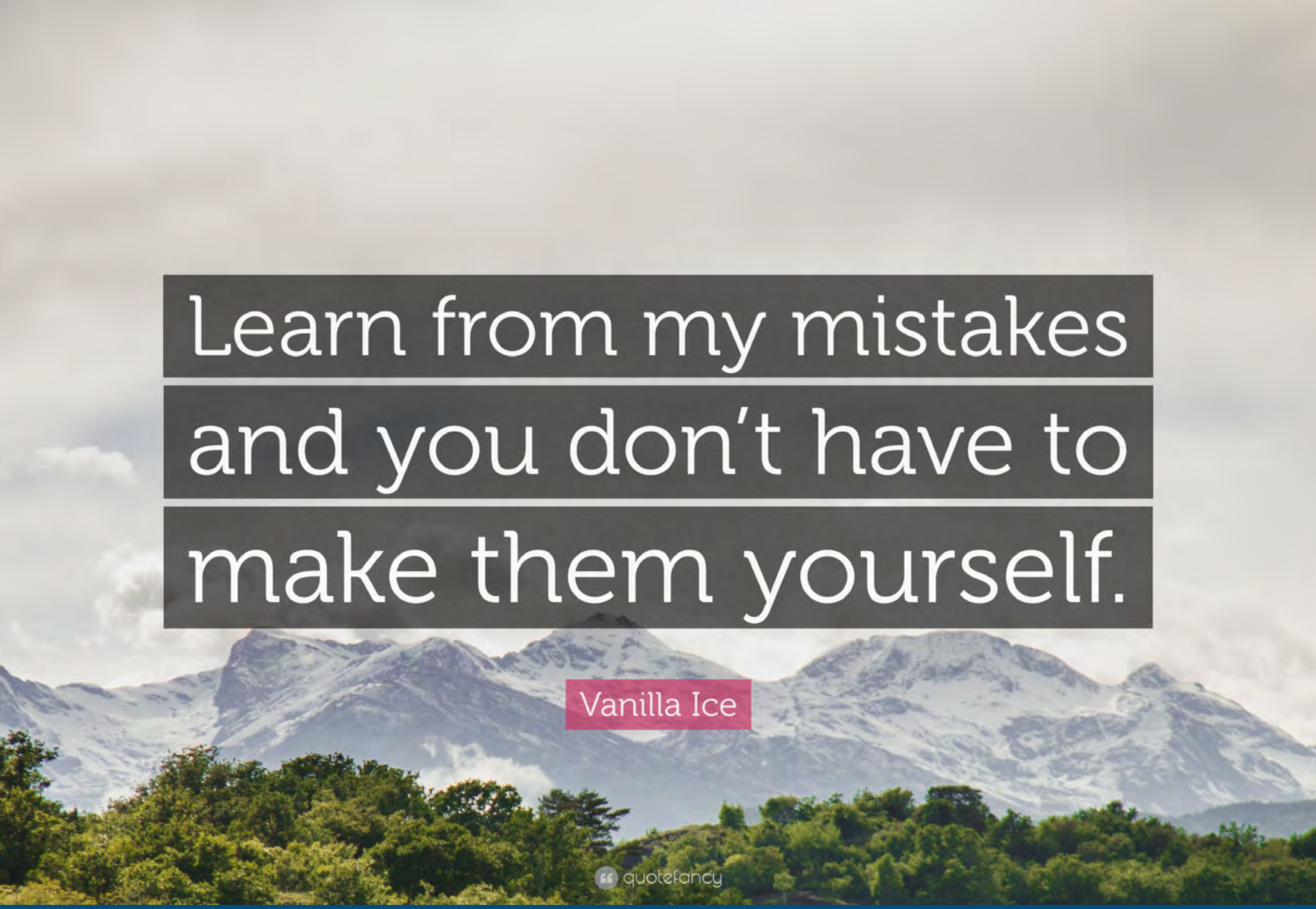


Technology that can help

- Rmarkdown
- Jupyter notebook
- Git
  - GitHub
  - Bitbucket
- Project management
  - Jira
  - Trello
  - Azure Boards

Peer review

- Data
- Model form
- Model build
- Validation
- Reasonability of results


Learn from my mistakes
and you don't have to
make them yourself.
Vanilla Ice
quotefancy

Questions about other things I've learned from my mistakes?

SOCIETY OF ACTUARIES®

talex.diede@milliman.com

somewhat different

Practical Aspects of Predictive Models

Session 35: Advanced Practitioners

Nick Hanewinckel, FSA, CERA
AVP and Actuary
SOA Predictive Analytics Symposium 2019/9/20

hannover re®

Disclaimer

The information provided in this presentation does in no way whatsoever constitute legal, accounting, tax or other professional advice.

While Hannover Rück SE has endeavoured to include in this presentation information it believes to be reliable, complete and up-to-date, the company does not make any representation or warranty, express or implied, as to the accuracy, completeness or updated status of such information.

Therefore, in no case whatsoever will Hannover Rück SE and its affiliated companies or directors, officers or employees be liable to anyone for any decision made or action taken in conjunction with the information in this presentation or for any related damages.

somewhat different

Bringing Along Non-Technical Stakeholders

hannover re®

A perfect Model?

Modellers may Accept a Model Using Several Statistical Tests

- Statistical tests, aside, stakeholders need to:
  - Understand where the model fits (Modelling Space)
  - Understand the data underlying the model
  - Become comfortable with variable impacts

- We will NOT cover technical model fit decisions

Source or Footnote

So Many Practical Considerations!

- As we go, I will point out other practical considerations as they come up

- Actuaries working in Predictive Modelling must be **vigilant** as new issues may arise

Finally

- We can also consider practical *applications* of predictive models!

somewhat different

Fit

hannover re®

Technical Expressions of Fit

What technical things get said?

- The RMSE is…
- The AUC is…
- The A/E is…
- R^2 is…

If the audience doesn't understand these…
will they trust the model?

What Could We Say Instead?

- We can define A/E (et al) by:
  - Subpopulations
  - Financial Measurements
  - Confidence Intervals
    - Uh oh…what do we mean exactly?

Give People What They Are Used To Seeing

- Example: many actuaries are familiar with Tabular Rates
  - We can express model predictions as such
- Actual-to-Expected ratios are also familiar
  - But…how to tell when deviations are just “noise?”

Attained.Age	AtoEcount_Testing	AtoEcount_Training	Exposure_Testing	Exposure_Training
55	1.013	1.076	109920.32	255020.18
56	0.995	0.983	107593.83	245303.27
57	1.100	0.961	99771.87	240546.97
58	1.060	0.990	100453.01	223251.52
59	0.953	0.994	92027.87	220556.96
60	0.974	0.994	94487.06	214252.16
61	1.007	0.956	95334.26	221700.08
62	1.013	1.031	86624.04	199585.80
63	1.053	1.029	68385.06	163853.79
64	1.082	1.007	67007.47	156172.64
65	1.020	1.032	66712.85	156428.86
66	1.012	1.032	65446.42	150424.29
67	1.033	0.991	56513.60	130692.79
68	0.964	1.022	49999.15	120226.50
69	1.039	1.007	47131.22	112353.83
70	0.967	1.096	49001.44	102214.57
71	1.035	0.971	41827.19	102298.81
72	0.916	0.965	43430.38	93754.59
73	0.975	0.990	41661.49	92141.64
74	1.009	1.009	37307.40	90117.15
75	1.033	0.979	34822.62	85351.35

Confidence Intervals

- Statisticians often mean X% confidence that a prediction lies within the interval
  - "Parameter Error" or "Model Error"
- Actuaries often mean X% confidence that the *result* lies within the interval
  - "Process Error"

Neither definition "nails" the concept of Prediction Error

Key Framework for Confidence Intervals

- Process Error – A “perfect” model still has residuals from the “natural process”
  - Example: A Poisson model expects 3.27 deaths (impossible).

- Parameter error – Errors in the model’s coefficients* or parameters
  - Example: A Linear Regression finds that β = 3, but expresses a 95% CI of [2,4]

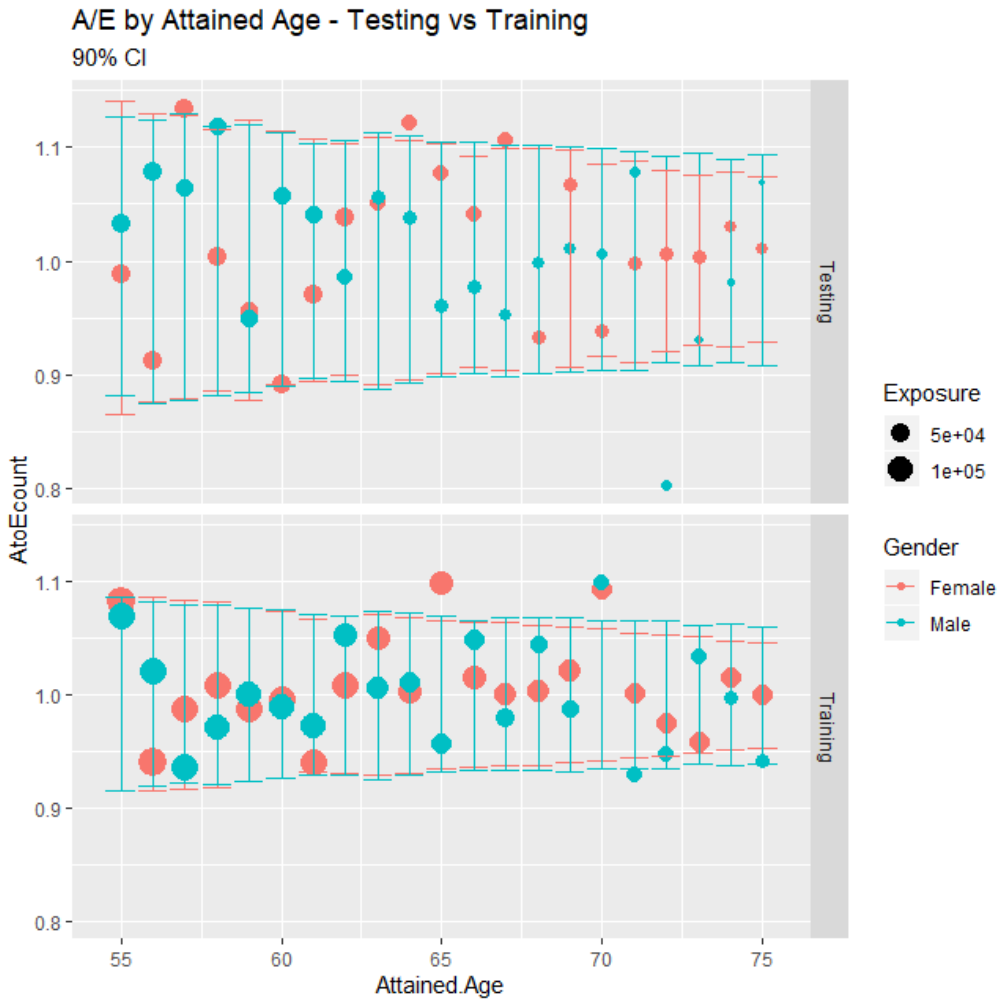
Thoughts – Which one is easier to measure? Why?

Practical Idea

- Express your results (e.g. A/E) with the CI for Process Variance
  - Communicate what this really means
  - How often are we outside this CI?

  The important point: this will not tell you the range of predictions, but rather the range of results if the model is correct behavior

Example – Process Error



- Poisson regressions have a clearly defined variance:
  - $\mu = \sigma$
- Therefore, we can generate poisson quantiles for the desired CI levels
  - in R:

```
qpois(CI, lambda=ExpectedDeaths,
lower.tail=TRUE/FALSE)
```

What does that tell us?

- What range of deaths are within X probability, assuming λ (qx) is correct?
  - More than x% out of CI → λ may be incorrect!

- It does NOT tell us the range of lambdas (qx) our model might produce
  - This is *parameter error*

somewhat different

Bonus Content

hannover re®

Examples

- Code and examples used in this presentation can be found here:
  - https://github.com/hanewinckel/PracticalAspectsPredictiveModels
- Bonus Content – An Additional Practical Consideration: Safe Splines?
  - Question: How can splines be unsafe?

somewhat different

Practical Applications *for* Predictive Models

hannover re®

Predictive Models Have Many Well-Known Uses

- Rate Setting – qx, disability rates, utilization rates, policyholder behavior
- Rapid Underwriting – lab-free scoring, risk classification models
- Marketing – Propensity to Buy

Other Practical Uses - Assumption Analysis

- This idea is an example of using a Predictive Model for Descriptive Statistics
  - Set up a model where existing assumption/basis is a “prior”
  - Where does your model deviate from experience?
  - What variables does this model find most meaningful?

- Thoughts – why might this be more meaningful than a simple A/E?

hannover re®

securian FINANCIAL®

Practical Aspects of Predictive Modeling

Lessons learned

Gary Hatfield
Sr Director and Actuary – Research and Analytics
Securian financial

9/20/2019

Mandatory Disclaimer

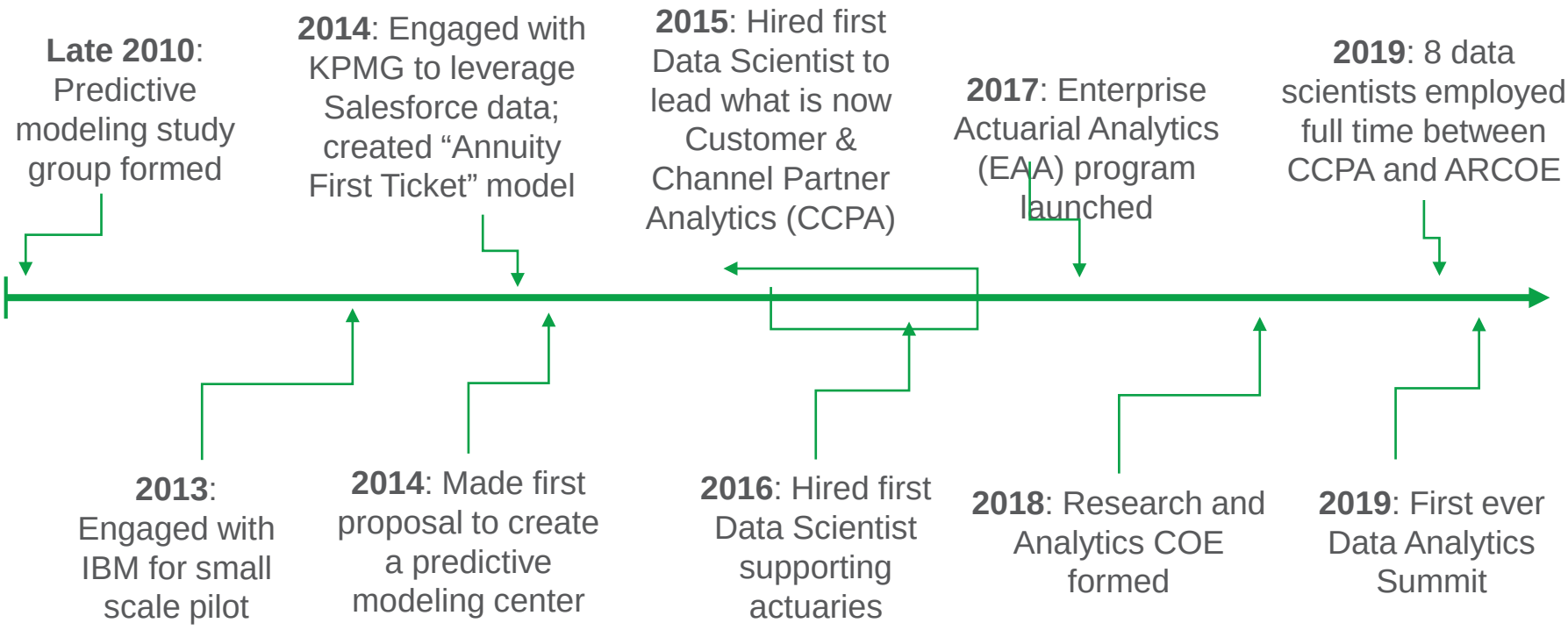
- Views expressed here are mine and mine alone. They should not be taken to represent those of Securian Financial Group.

Agenda

- A bit about Data Analytics at Securian
- Six Lessons Learned

Data Analytics at Securian

Timeline for Data Analytics at Securian



What is EAA about?

Purpose

Demonstrably enhance the capabilities of our actuarial functions to perform analytics

Objectives

- Grow analytical expertise
- Improve the quality, completeness, accessibility, and understanding of data
- Provide tools and processes to effectively conduct actuarial analysis

Structure

- Use Case Work Stream
- Foundation Capability Work Stream

EAA Project Topics

- Accelerated Underwriting
- PRT mortality
- Group pricing loads
- Annuity lapse model
- Advisor behavior

Six Lessons Learned

Lessons

1. Prioritize projects that business leaders are fired up about
2. Hire mix of actuaries and data scientists
3. Often, the low hanging fruit is just gathering data and doing EDA
4. Engage with data engineers
5. Partner with law, privacy, and compliance
6. As much as possible, shepherd to production

Thank You!

Insurance products are issued by Minnesota Life Insurance Company or Securian Life Insurance Company, a New York authorized insurer. Minnesota Life is not an authorized New York insurer and does not do insurance business in New York. Both companies are headquartered in St. Paul, MN. Product availability and features may vary by state. Each insurer is solely responsible for the financial obligations under the policies or contracts it issues.

Securian Financial is the marketing name for Securian Financial Group, Inc., and its affiliates. Minnesota Life Insurance Company and Securian Life Insurance Company are affiliates of Securian Financial Group, Inc.

Securian Financial is the marketing name for Securian Financial Group, Inc., and its affiliates.

Securian Financial Group, Inc.
securian.com

400 Robert Street North, St. Paul, MN 55101-2098

F00000 Rev 1-2019 DOFU 1-2019
720672