Session 73L, Advanced Analytics and Predictive Modeling in Loss Reserving

Presenters:

Mark M. Zanecki, ASA, MAAA

S

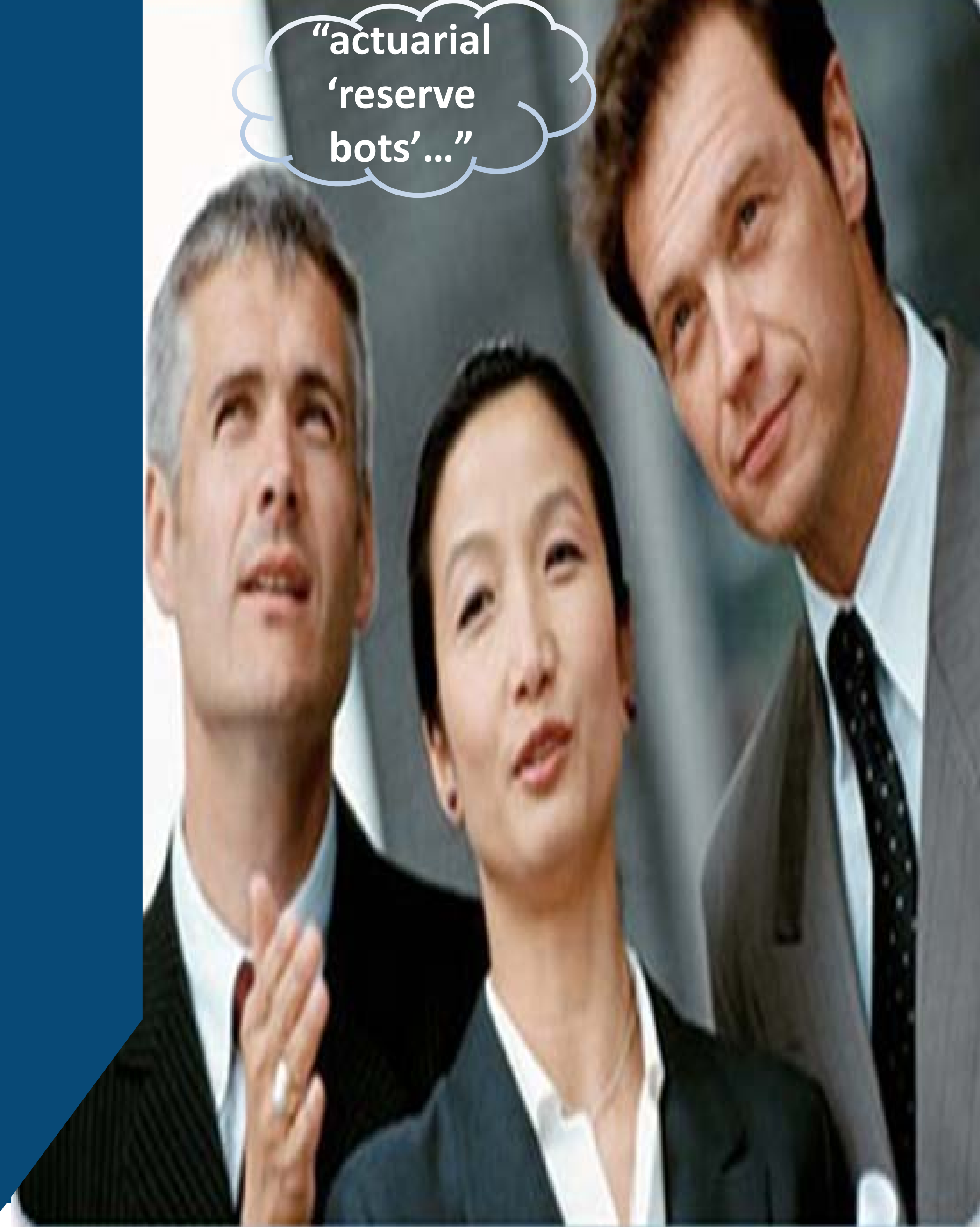
Session: 73

Advanced Analytics and Predictive Modeling in Loss Reserving – First Generation Machine Learning

Author:
Mark Zanecki ASA, MAAA

SOA Antitrust Disclaimer
SOA Presentation Disclaimer

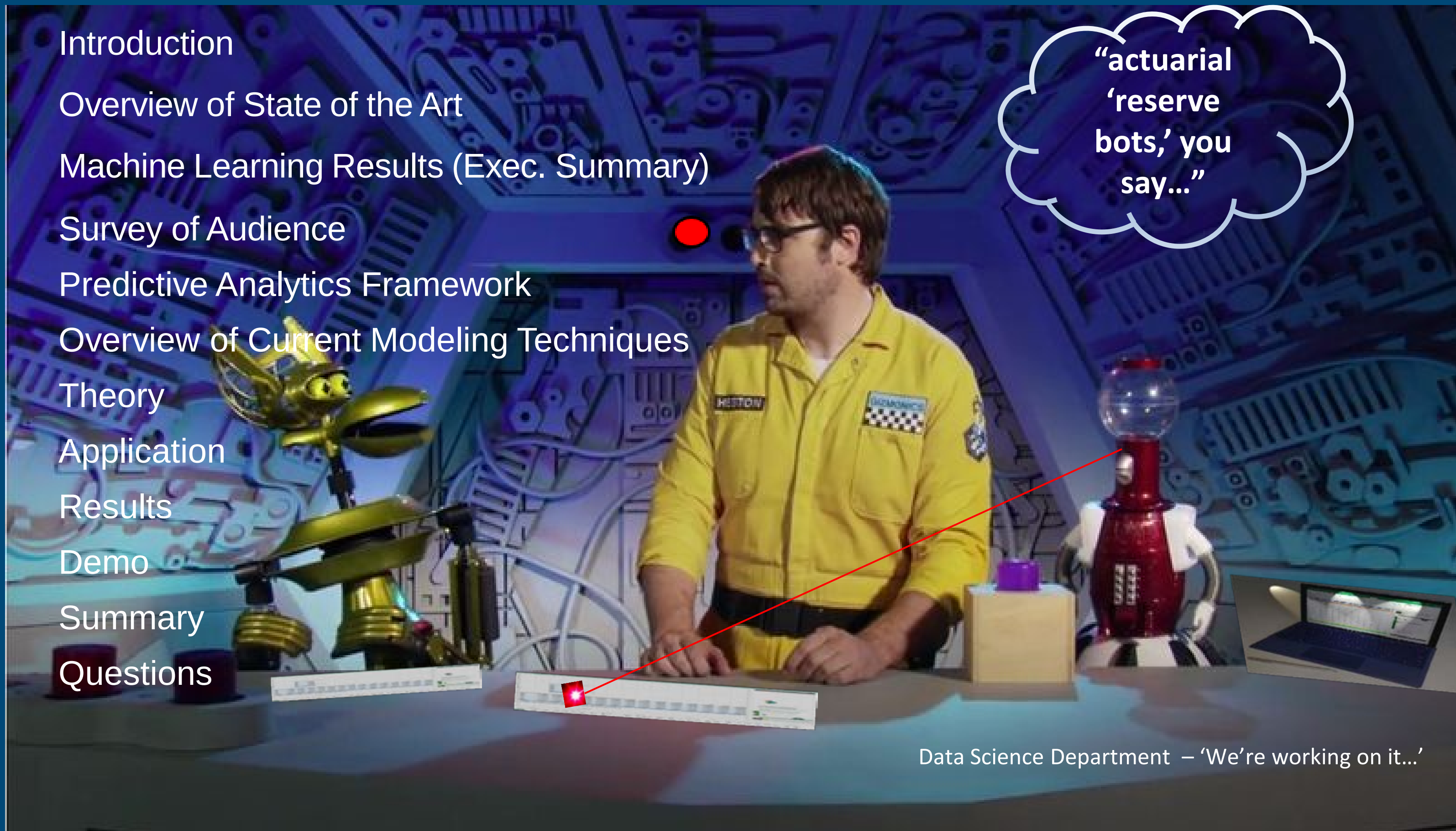
2018 SOA Health Seminar
Austin, Texas
June 25-27



IHA Consultants Inc.

Agenda

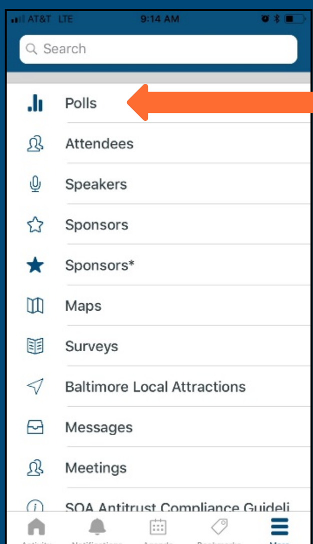
- Introduction
- Overview of State of the Art
- Machine Learning Results (Exec. Summary)
- Survey of Audience
- Predictive Analytics Framework
- Overview of Current Modeling Techniques
- Theory
- Application
- Results
- Demo
- Summary
- Questions



Credit: IHA Consultants Inc. Credit: Microsoft Credit: Mystery Science 3000

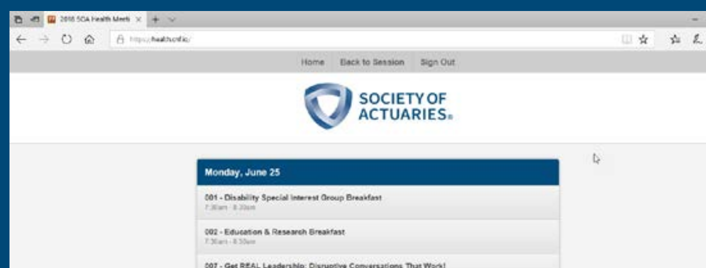
To Participate, look for Polls in the SOA Event App or visit **health.cnf.io** in your browser

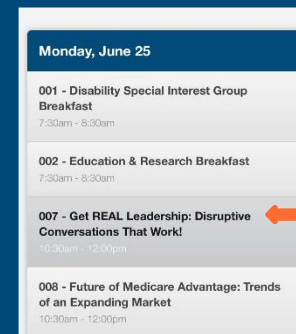
Find The Polls Feature Under **More** In The Event App



or

Type **health.cnf.io** In Your Browser





Choose your session

Live Content Slide

When playing as a slideshow, this slide will display live content

Social Q&A
(poll closes after advance past slide)

Live Content Slide

When playing as a slideshow, this slide will display live content

Poll: (warm-up question#1)In machine learning / predictive modeling, is there a distinction between prediction and forecasting?

Live Content Slide

When playing as a slideshow, this slide will display live content

Poll: (warm-up question#2)Which type of prediction would you expect to be more difficult?

Live Content Slide

When playing as a slideshow, this slide will display live content

Poll: What is your interest in attending session 73?

Live Content Slide

When playing as a slideshow, this slide will display live content

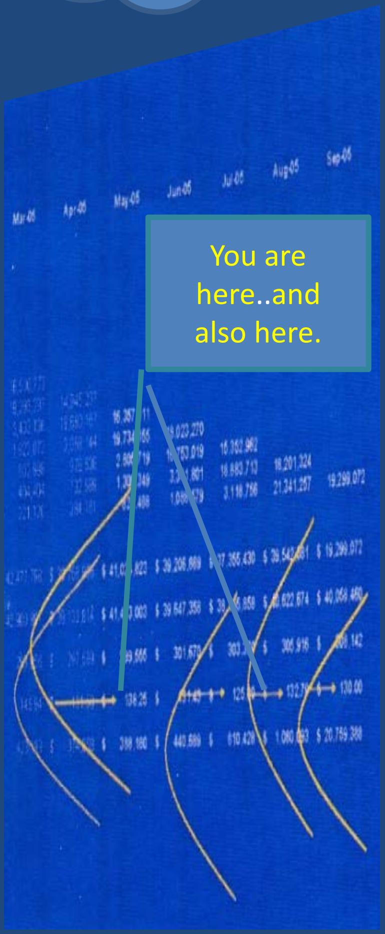
Poll: Please self-rate your background / knowledge with respect to machine learning/predictive modeling theory.

Do you know where you are in your distribution?

At 25,000 foot Level – we have ‘CVM’:

There are three modeling camps:

- **“(Cents) -Centralists”** – filter(smooth) results to remove fluctuations and rely on central limit theorem (CLT). Point estimate technique. Actuaries and econometricians favor this camp.
- **“(Vols) - volatility”** – model volatility (tame volatility and results follow by CLT in distribution form. Some actuaries , a good number of econometricians, many statisticians and every investment bank stock options quant.
- **“(MLs) -Machine Learning”** – using historical data develop models using as many predictor variables (called features) as required combined with varying correlation and partitioning methodologies. Membership is open if you know computer algorithms, computer technology, data considerations, statistics, mathematics, econometrics and enjoy “challenges.” Caution: Finding local optima and not global optima is not always acceptable nor sufficient. (‘no second best suffices.’)



Theory, Application and Results Learned (Exec. Summary):

Theory:

- Regression defines error as:

  error = realized value – predicted value(via model)

- Machine Learning partitions error as:

  $$error = bias + variance$$

- You can do better, using the following extension:

  $$error = bias + stable\ variance +$$

Application:

- The ideal machine learning "factory" relies on automation in the following sense(s):
  a) continuous re-training of existing models on new data,
  b) has the capability of creating new models when old models are insufficient (also part of the automation process.) The challenge is to create this capability.
- Successful machine learning automation frameworks rely on many layers beginning with data and proceeding to many layers of modeling.
- The ability to detect and quickly recover from algorithmic error or data error is critical.
- Machine learning requires high compute speed, vast memory size, very high storage size (Tera/Peta scale) which is delivered via gpu compute environment at the high end and via Hadoop type at the lower end.
- Driving firm value from data science requires more than just staffing a department, purchasing off-the-shelf-software and running large amounts of data thru models.

Theory, Application and Results Learned (Exec. Summary continued):

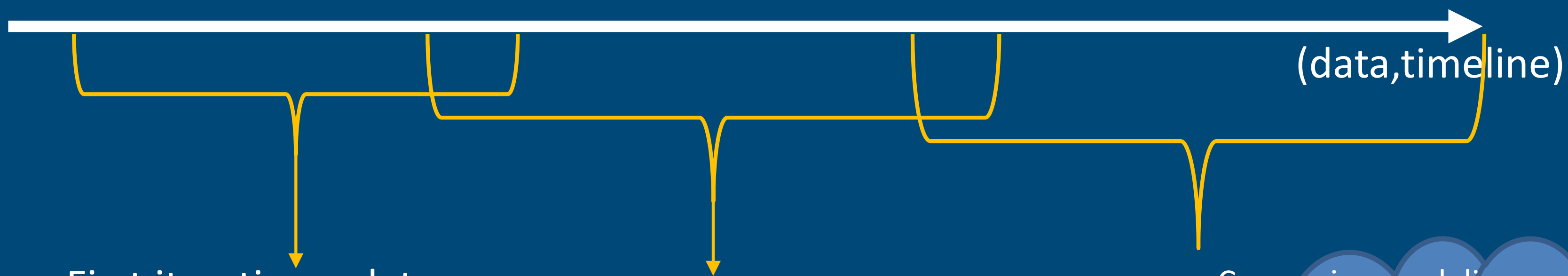
Result(s):

- Saving the Best for Last – US Stock Price Prediction Analysis: Using machine learning based upon gpu servers; 8,500+ stock symbols are scanned and next day trading range is predicted with accuracy greater than based upon prior 20 trade day analysis.
- Is there a similar result for loss reserves? Ans: Yes, and it is
- "No risk = no reward was verified." The best stock(s) reported alpha at no greater than zero, none reported alpha > 0 and majority had alpha < 0, for short duration horizons. Upward trend /downward trend was attributable to and
- Implication: The "winner" will be those firms that can measure, model and price/package volatility (e.g. embrace it) – the future for pure smoothing is limited by comparison.

Announcement(s):

- The audience is encouraged to use the interactive session evaluation app.
- There is finally a useful app for my Windows Smartphone (stock price prediction app) that actually uses Mobile Excel. Who would have expected to have to write it yourself?

Why do you need a Machine Learning Factory (automation)?



First iteration – data, model estimation

- model1
- model2
- model3
- …..
- model(n)

Second iteration – data, model estimation

- model1
- model2
- model3
- …..
- model(n)

Comparing modeling epochs:

1) can have same model with different parameters

OR

2) new model with different parameters.

Modeling **must check all models** and then via automation pick optimal model.

You need GPU server(s) to do this…..

Predictive Analytics / Machine Learning - Audience Survey

TBD – show of hands.

Live Content Slide

When playing as a slideshow, this slide will display live content

Poll: What kind of a session would interest you more?

Live Content Slide

When playing as a slideshow, this slide will display live content

Poll: How would you rate your employer's interest in machine learning/predictive modeling?

Live Content Slide

When playing as a slideshow, this slide will display live content

Poll: Please choose from the following(chose as many as applicable).

Predictive Analytics / Machine Learning Framework:

There are **three levels** of predictive modeling.

Level 1:
Analytical methods (descriptive statistics) that are based on summarizing historical data stored in data lake / date warehouse. Typically use Tableau for visualization in combination with a business intelligence tool.

Level 2:
Predictive Modeling which focuses on identifying, classifying and quantifying various relationships learned from past data using supervised learning and unsupervised learning techniques. Model(s) are then used to predict various outcomes. Base model is some form of logistic modeling. More advanced forms include ensemble modeling (bias reduction), decision trees, gradient boosted methods (GBM), structured vector machines (SVM), and numerous forms of artificial neural net models (ANN).

Level 3:
Prescriptive Modeling which focuses on first identifying members in a population based on predictive modeling results and then forming an action or recommendation as a second step.

Overview of Current Modeling Techniques (next 8 slides):

Reserving as Predictive Problem

vs.

Regression as Machine Learning Problem

- Model Considerations
- Data Considerations
- Assumption Considerations
- SLA - Time to Complete

CY	Dur	Incremental Loss or Cumulative loss
2014	0	6,000
2014	1	7,000
2014	2	5,000
2015	0	7,000
…	…	…
2017	1	???
2018	0	???

Loss Reserve as Linear Regression Problem

X		y
CY	**Dur**	**Cumulative loss**
2014	0	6,000
2014	1	7,000
2014	2	5,000
2015	0	7,000
…	…	…
2017	1	???
2018	0	???

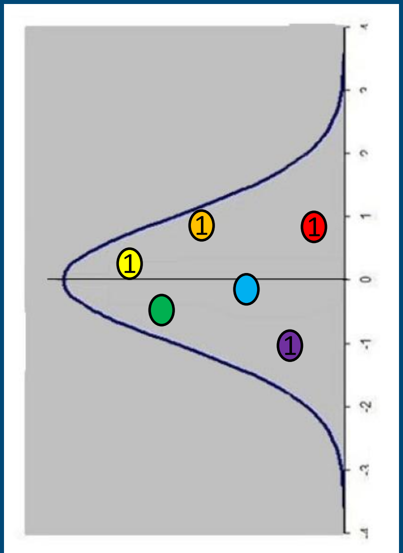
The Regression Problem

$Solve\ \mathbf{Xb} = \mathbf{y}$

where X is design matrix
y is response vector variable

- Requires error assumption – Normal is typical.
- Assumes homoscedastic error, independence of columns and other assumptions. (CF violates this assumption. Determinant is close to 0.0)
- Using linear predictors in non-linear situations has unacceptable error rate.

Stochastic Time Series



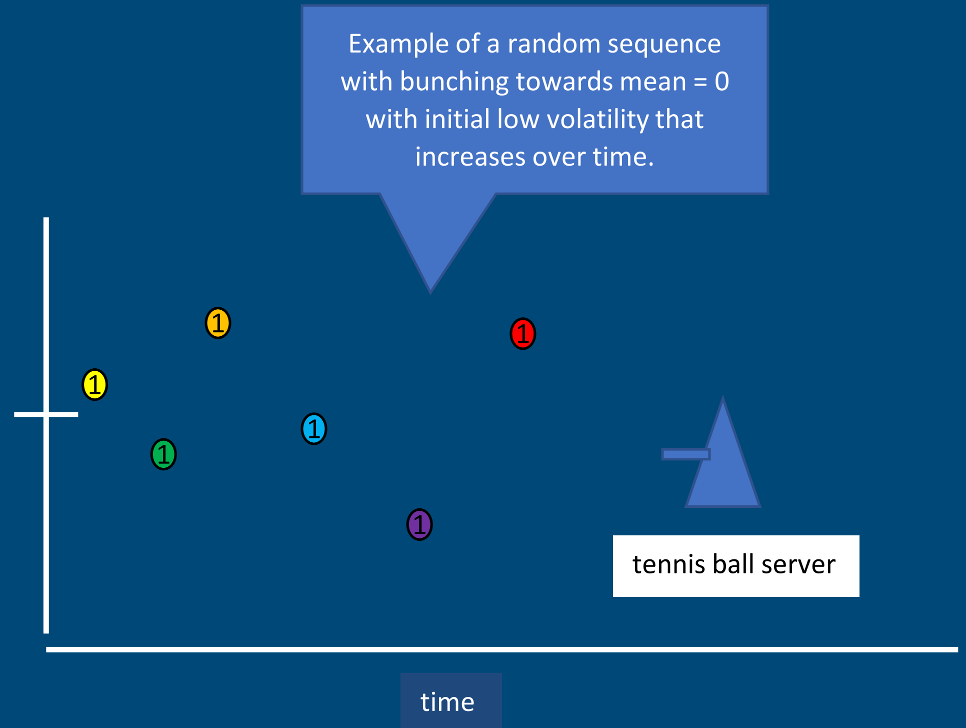
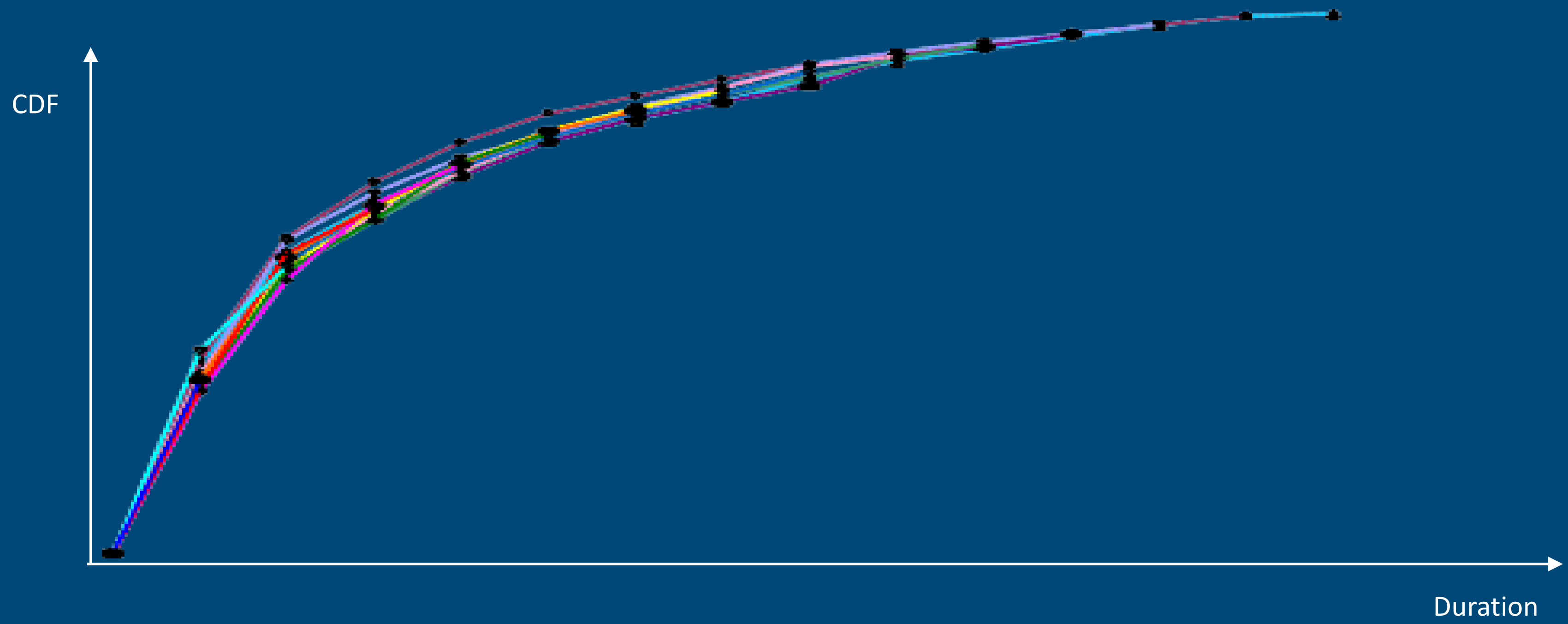


Figure - as viewed from wall and overhead cameras

Illustrative models include:

- Moving Average
- Auto-regressive
- ARIMA – combination of moving average & auto-regressive
- More sophisticated: GARCH which involves modeling time varying volatility (various flavors). This requires solving Maximizes Log of Likelihood function which is a non linear optimization problem. Mean-reversion processes only.

Simulation



- Simulation is useful to estimate the predicted distribution of values generated by numerous path iterations.
- In life insurance reserving, simulation is the tool for valuation for secondary guarantee reserves.
- Healthcare loss reserves use simulation to develop best estimate or to estimate variability of best estimate.

Generalized Linear Model (GLM)

(Predictors) (Response)

CY	Dur	Incremental Loss
2014	0	6,000
2014	1	7,000
2014	2	5,000
2015	0	7,000
…	…	…
2017	1	???
2018	0	???

$$Y_{CY, Dur} = \exp(\beta_0 + \beta_{CY} + \beta_{Dur}) + \varepsilon$$

Log link function

Linear combination of explanatory variables predicts incremental losses, based on CY and Dur and other identified predictors.

	Dur						
CY	0	1	2	3	4	5	6
2012	80	130	123.5	135	150	200	222
2013	110	121	115	140	155	210	
2014	105	116	110	138	160		
2015	90	120	114	125			
2016	96	106	100				
2017	120	132					
2018	122						

Comment(s):

- Error distribution: Tweedie/Gamma/Poisson
- Predictors: Calendar year, duration/development period, etc.
- Results sensitive to error assumption.
- Available in Python, R, Matlab, SAS and other packages

Theory: Machine Learning (Supervised)

No limit to number of features. Discovery of smallest, reliable feature set is the problem.

"Train model" using historic data features and apply to future data predictors.

To each Y_i associate a function of X variables. (This is trial and error process.)

[X] [y]

CY	Dur	Var_1	Var_2	Var_3	…	Incremental Loss or Cumulative loss
2014	0					6,000
2014	1					7,000
2014	2					5,000
2015	0					7,000
…	…					…
2017	1					???
2018	0					???

New Predictors

Theory: Machine Learning (Supervised) – Thought Experiment(s)

As X variables vary, Y_i varies – static model learned from historic data for that period.

What happens if over time, Y_i varies not only with X variables but also by passage of time? Ans: frequemt re-training. (Recurrent Neural Net is typically used.)

What is one of the fundamental differences between machine learning vs classical regression? Ans: From linear algebra the column space must be full span, non-zero determinant, independent and ideally non correlated. Not forgetting heteroscedasticity.

(X) (y)

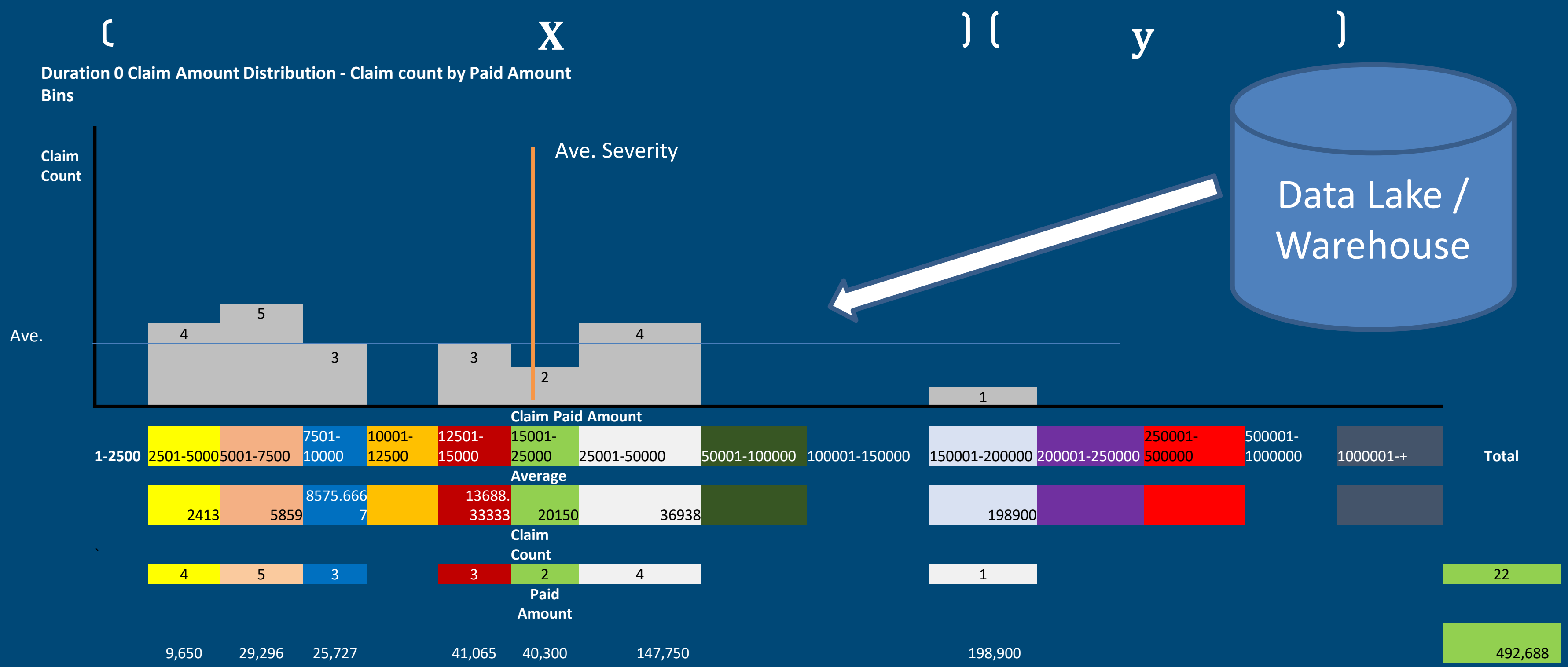
CY	Dur	Var_1	Var_2	Var_3	…	Incremental Loss or Cumulative loss
2014	0					6,000
2014	1					7,000
2014	2					5,000
2015	0					7,000
…	…					…
2017	1					???
2018	0					???

New Predictors

Machine Learning (Supervised) – Thought Experiment

Consider for claim type (hospital inpat., outpatient; physician inpat., outpat, office , drug etc.) modeling by duration using claim amount bins and claim count.

What are the pros & cons? Ans: Reduces problem to Poisson Count (by claim detail type and by bin) but with population origination, timing and trend considerations.



New Predictor Model – Bins(Ave Severity, Claim Count) by Duration Claim type with detail available from data warehouse

Claim level models allow us to understand *why* development is changing

What we need is proper level of claim detail and proper level of predictive modeling. (credit: “captain obvious”)

{ Level }	{ Predictors }	{ Response }	{ Comment:}
Model is a function of data features and assumptions.	Fewer the better	Credibility, variability, consistency and accuracy issues	Modeling and data decisions are interdependent.
Claim Aggregate	Paid Amounts by claim type	Relies on assumption that past predicts future	Too high a level – too much detail lost (is the word on the street)
Claim Aggregate with Additional Disparate Data	Paid Amounts by claim type plus other predictors	Correlation consistency over time	This is more art than science
Claim Detail	Claim type, Dx(s), Procedure Codes, Provider, Age, Gender, Plan Code,… etc.	Variance ratchets up	Too fine a level for analysis – great for descriptive
Claim Detail Disparate Data	Claim type, Dx(s), Procedure Codes, Provider, Age, Gender, Plan Code,… plus other predictors	Variance ratchets up with a side order of correlation consistency and credibility	Big Data Baby! How useful has yet to be determined.

Application: Observation(s):

Claim Analytics:

The data lake/data warehouse supports all levels of detail for claim, premium-billing and provider. Claim level analytics can be automated into dashboards via BI reporting packages. It is assumed that claim reserves are developed by claim type, outliers are removed and adherence to SOA Health Valuation ASOP/Manuals . Value-add, for analytics ,is knowledge of the totality of claims with details and how similar for dissimilar your sub-population. This is key for value based contracting for ACO and Medicare.

Predictive Modeling:

We can choose to model at any level which will give acceptable results. Let's agree to chose an overall approach using aggregate claim data(top-down) which is supplemented with a bin(ave. severity, freq.) by duration (if we need it). At all times we can drill further to lowest detail.

Top-Down approach:

At the highest level, the effect of all the variables that can impact loss reserve modeling is captured sufficiently and measured in (mean, variance) space. (Similarly for stock price prediction only there are more variables and "'Lucas, Muth, Sargents' (famous macro-economics paper) irrational expectations forecasting effect.)(see https://en.wikipedia.org/wiki/Rational_expectations)

In-the-Middle approach:

Partition aggregation data by duration into bins(ave. severity, freq.). The associated claim detail predictor variables are still attached if we need an additional level of predictive modeling and / or support reporting at any level of descriptive analytics.

Bottom-up approach:

At the lowest level, we can train models (supervised training) using historic data (big data) and predict future variables of interest. Aggregate up to measure impact in (mean, variance) space. Danger(s) include:

1. Over-fitting. Over fitting means can predict the past with high accuracy but can not predict future with sufficient accuracy.
2. Finding non stable local optima result that vanishes with new data.
3. Identifying new correlated predictors, that are not casual in nature or have inconsistent, varying correlation.

Digression – Bounded Cauchy Sequences of Real Numbers

Every Cauchy sequence of real numbers is bounded.

A sufficient condition is that at high enough index (n) the difference between consecutive terms approaches, in the limit, 0.

MAIN POINT is WHAT MEASURE was used to arrive at the conclusion:
Possible measures:

- original individual value(s) approach M
- first difference between contiguous values
- second difference of the first difference
- trend between contiguous values
- moving average (length=k)
- variance of moving average (length = k)

Any of these measures would prove the same conclusion, just not as succinctly.

Example:

CF values	0.89964	0.90909	0.91496	0.92142	0.937282	0.95323	0.974453
first difference		0.00944	0.00587	0.00646	0.015853	0.01595	0.021216
second difference			-0.00357	0.00058	0.00939	0.00010	0.005261
average-drop-high-drop-low							0.927201
average							0.930014
variance							0.000705
trend		1.01050	1.00646	1.007064	1.017205	1.01702	1.022257
average trend							1.013419
variance of trend							4.05E-05
Trended Completion Factor			0.927201	1.013419			0.939643
Trended with Variance Completion Factor			0.927201	1.019786			0.945547

For aggregate data, the actuary is afraid of conflating trend with variance, so to be conservative uses a smoothing which eliminates any adjustment.

Example (continued):

CF values	0.89964	0.90909	0.91496	0.92142	0.937282	0.95323	0.974453
first difference		0.00944	0.00587	0.00646	0.015853	0.01595	0.021216
second difference			-0.00357	0.00058	0.00939	0.00010	0.005261
average-drop-high-drop-low							0.927201
average							0.930014
variance							0.000705
trend		1.01050	1.00646	1.007064	1.017205	1.01702	1.022257
average trend							1.013419
variance of trend							4.05E-05
Trended Completion Factor			0.927201	1.013419			0.939643
Trended with Variance Completion Factor			0.927201	1.019786			0.945547

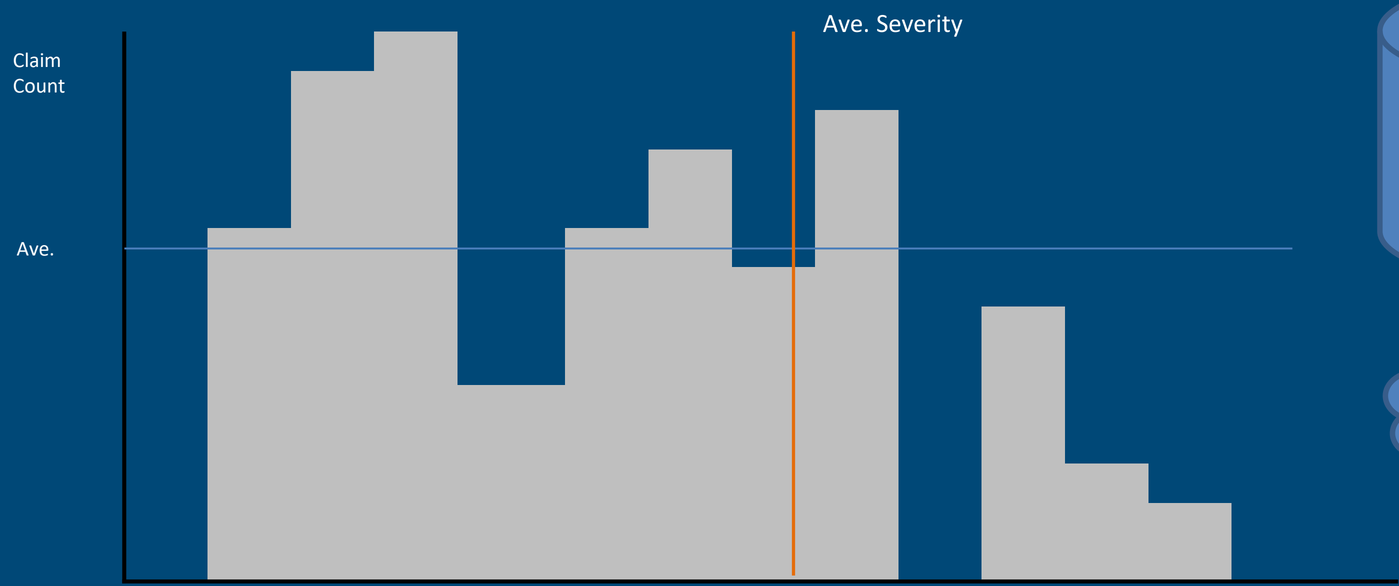
If we were to examine using bin model(ave. severity, freq.), what would we discover?
Ans. That different bins now have different ave. severity and different claim counts. Drilling deeper can report by claim Dx, Proc(x), Plan Code etc.
Should we use this new information?
Maybe..

For aggregate data, the actuary is afraid of conflating trend with variance, so to be conservative, uses a smoothing which eliminates any adjustment.

High Duration Cumulative Distribution

Consider for claim type (hospital inpat., outpatient; physician inpat., outpat, office , drug etc.) modeling by duration using claim amount bins and claim count.

At high duration, small claim amount bins do not change much and we see 'small' changes in mid to high claim bin amounts. The bin distribution is 'essentially stable', the average severity and average claim count by bin change in small increments which can be treated as <u>variance process only or combined trend and variance process</u>. This is the basis of completion factor methodology.

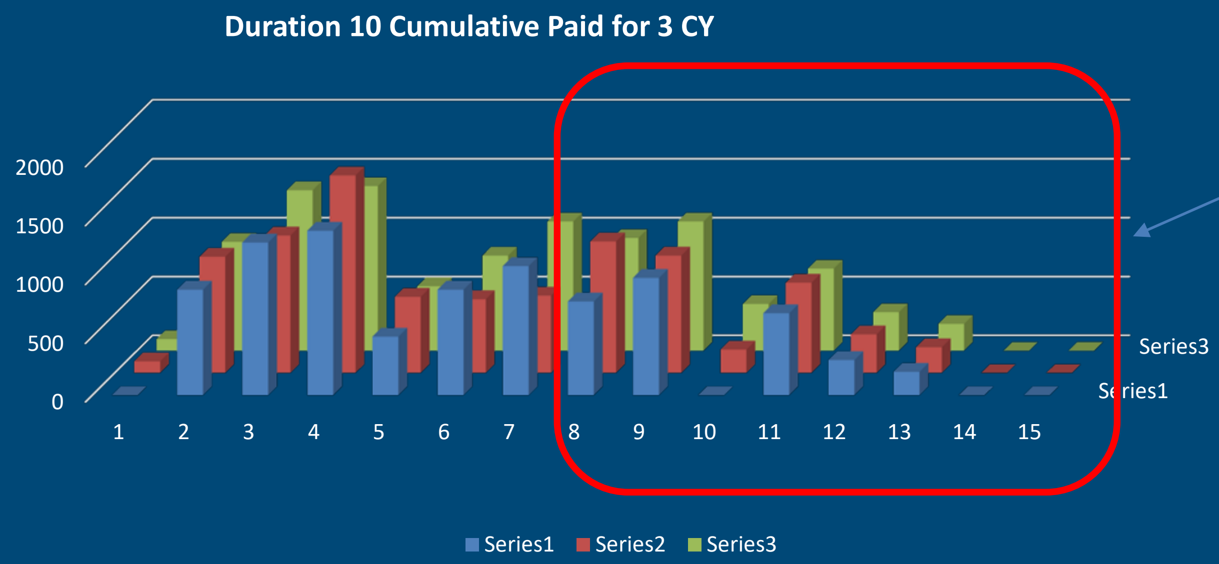


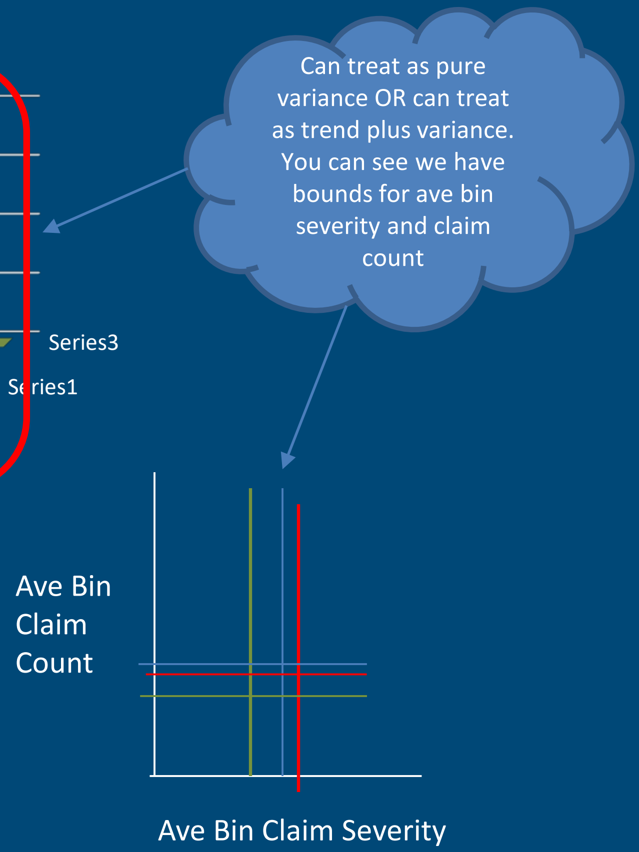
Claim Paid Amount	2501-5000	5001-7500	7501-10000	10001-12500	12501-15000	15001-25000	25001-50000	50001-100000	100001-150000	150001-200000	200001-250000	250001-500000	500001-1000000	1000001-+	Total
Average	2413	5859	8576	9575	13688	20150	36938	56000		198900	210000	35000			
Claim Count	900	1300	1400	500	900	1100	800	1000		700	300	200			9100
Paid Amount	2,172	29,296	25,727	4788	41,065	40,300	29550	56000		198,900					427,797,600

High Duration Cumulative Distribution

Consider for claim type (hospital inpat., outpatient; physician inpat., outpat, office , drug etc.) modeling by duration using claim amount bins and claim count.

At high duration, small claim amount bins do not change much and we see ‘small’ changes in mid to high claim bin amounts(tail). The bin distribution is essentially stable, the average severity and average claim count by bin change in small increments i.e. a variance process only. This is the basis of completion factor methodology.

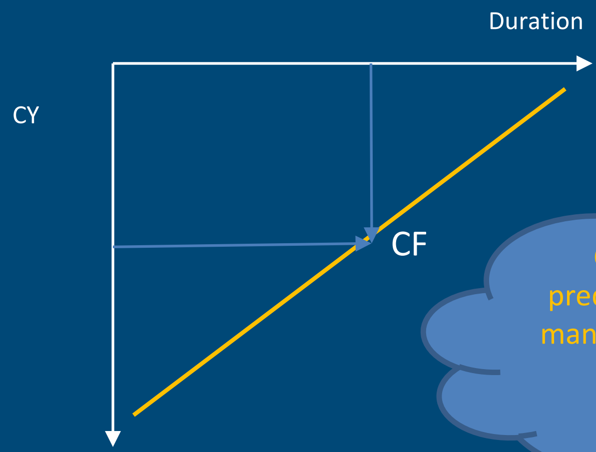




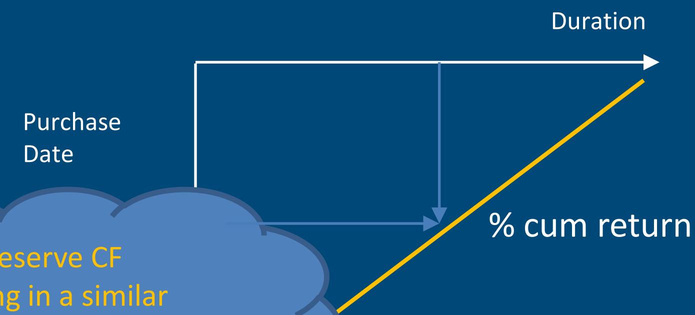
- Can we create a predictive model for the incremental amount paid in duration 10 using detail claim data? Ans: Yes
- **Can we create a predictive model using (mean, variance) using bins without the need of claim detail? Ans: Yes**
- **Conceptually, is the result of predictive bin model similar to predictive model using aggregate data? Ans: Yes**

Cumulative Return for a Stock has Similar Upper Triangle Analysis

Reserve Cumulative CF by CY



Cumulative Return for Stock by Purchase Date



Can we treat Reserve CF predictive modeling in a similar manner to predicting stock price return modeling?
Ans: Yes

- Completion factor varies over duration.
- Completion factor varies by CY
- There are fewer variables that affect CF as compared to stock. Health variables are somewhat controllable – just program the claim box.

- Stock price varies over duration.
- Stock price varies by purchase date.
- Many, many variables affect stock price, non are controllable and some difficult to measure.

Top-Down: Predictive model for (mean, variance)
Or
Bottom-up: Predictive model using vector of predictor variables to obtain increment then calculate (mean). No variance.

Markov-Chain Monte Carlo approach used to produce a simulated range around *GLM-based* deterministic estimate

Individual Claim Level Reserving / Aggregate Claim Level Reserving: Data Dictionary for Predicting Claim Incidence (Logistic Model) and Claim Severity for Bin Frequency, Severity by Duration (which supports full analytics) and Traditional Aggregate Healthcare Reserve Modeling

Insured Group Characteristics
- Policy start/ end date
- Months with company
- Number of employees
- Health Benefit Plan Code
- Rating Class
- Group Code
- Industry Code
- Funding Arrangement

Claimant Characteristics
- Age
- Gender
- Previous Medical Event- Dx
- Marital status
- Years employed
- Job level
- HSA Election Amounts
- Sick Days

Premium-Billing
- Payment history
- Payment plan

Loss Characteristics
- Procedure Codes
- Diagnosis: Dx1, Dx2
- Shot term Disability Claim
- WC Claim
- Allowed Amount
- ER Claim?
- State & Zip code
- Claim Receipt Date
- Date of Service

Health Providers
- Medical service provider(s)
- Pharmaceuticals
- PBM(s)

Geo-Location
- MSA- Urban/Rural
- Distance to work

Claim Type

Outlier Claim

Utilizing more data allows for increased understanding of losses

IHA Consultants Inc.

Data Elements for Predictive Stock Price Model (Fintech)

[X] [y]

Stock price(t-1)
Dividend
p/e
Volume(t-1)
Volume(t-2)
Industry
Market cap
News(t-1)
News(t-2)
Sector
Market Basket
S&P 500 index
Number of internet searches
Patent / copyright filing
Competitors stock price(t-1)
Earnings report

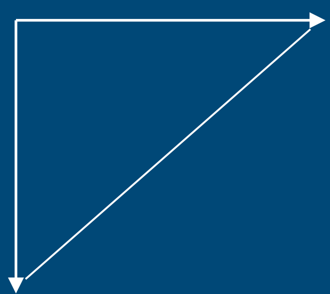
Where:

X: is matrix of predictor variables
Y: close stock price vector

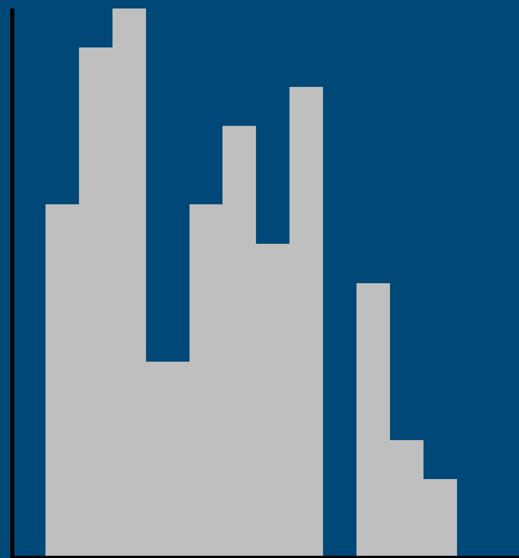
Results: Predictive Modeling Data Levels

Top-Down

Aggregate Data

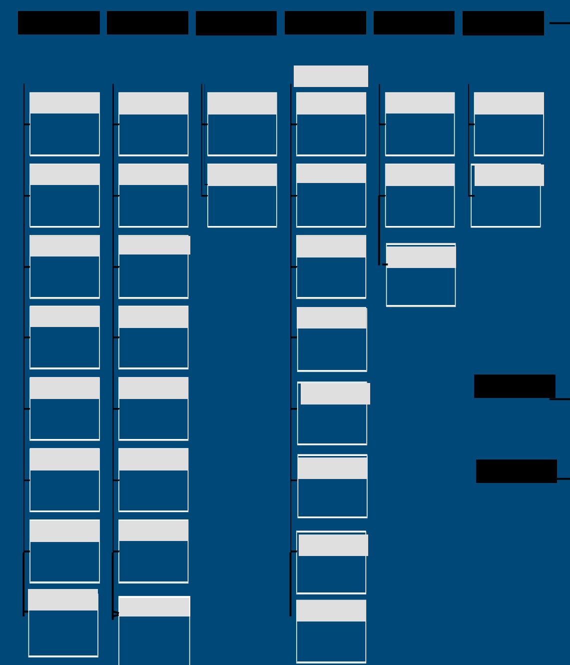


Middle-of-Road



Claim Paid Amount	2501-5000	5001-7500	7501-10000	10001-12500	12501-15000	15001-25000	25001-50000	50001-100000	100001-150000	150001-200000	200001-250000	250001-500000	500001-1000000	1000001-+	Total
Average	2413	5859	8576	9575	13688	20150	36938	56000		198900	210000	35000			
Claim Count	900	1300	1400	500	900	1100	800	1000		700	300	200			9100
Paid Amount	2,172	29,296	25,727	4788	41,065	40,300	29550	56000		198,900					427,797,600

Bottom-Up



Predictive Modeling Levels

Top:
- (mean, variance)
- Claim Type
- Aggregate Data

Middle:
- (mean, variance)
- Claim Type
- Bin(ave. severity, freq.)

Bottom:
- Claim amount or claim amount increment and claim counts
- Claim type
- Detail claim predictor variables

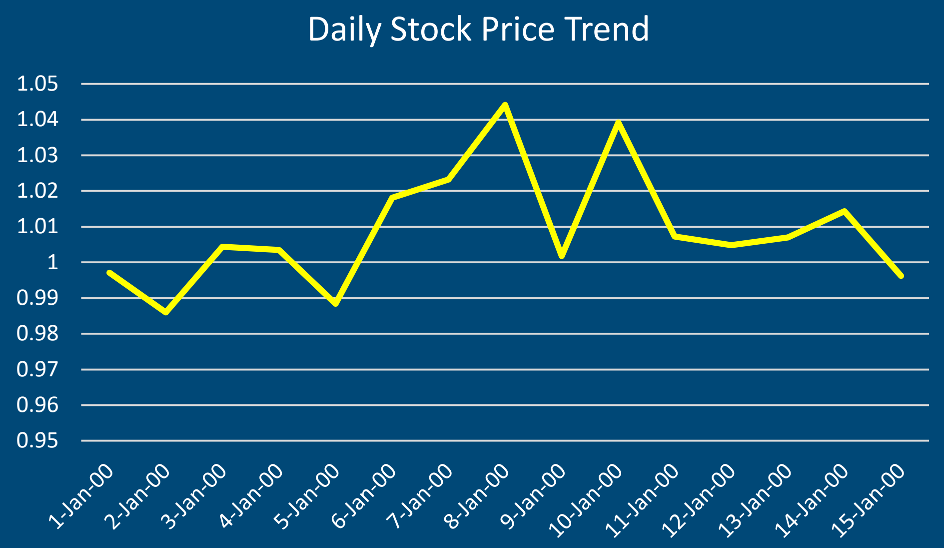
Predictive Use:

Top:
- Main model
- Fast screening model

Middle:
- Main model
- Confirming model
- Main model (reinsurance)

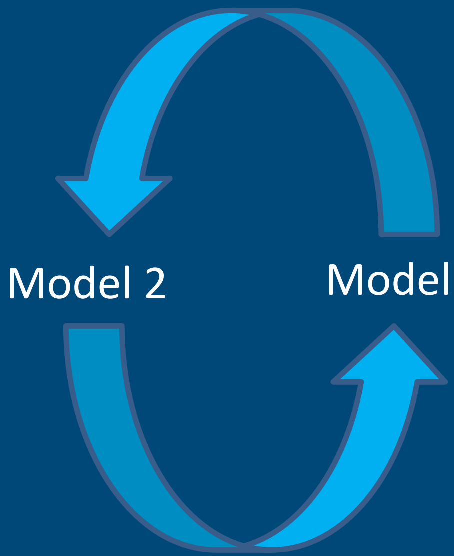
Bottom:
- Main model (reinsurance)
- Confirming model

Predictive Stock Price Modeling – Lessons Learned (Fintech)



Comment(s):

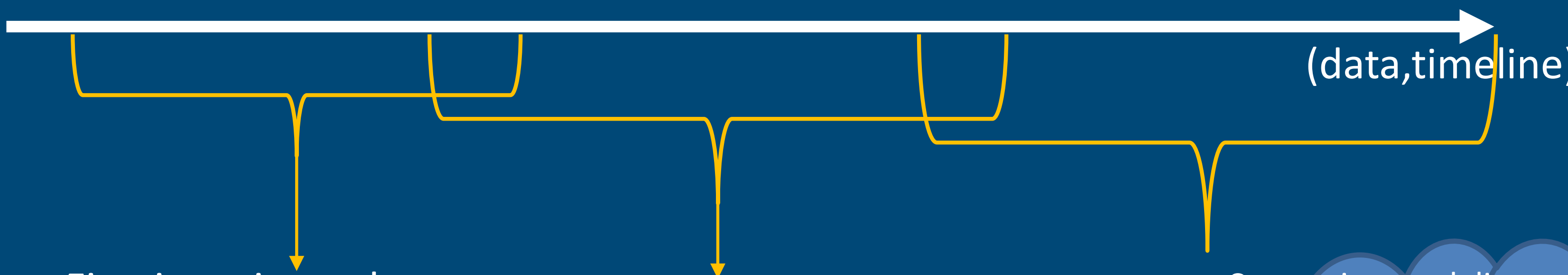
- Clearly have trend and variance present.
- Other than CF cap at 1.0, stock price series is 'similar' (albeit more volatility) than CF series, so conceptual similar modeling, recognizing completely different predictor variable sets and propensity effects.



Comment(s):

- Predictive modeling of stock price process requires more than 1 predictive model.(Many layers.)
- Predictive modeling factory automation is required.

Why do you need a Machine Learning Factory (automation)?



First iteration – data, model estimation

- model1
- model2
- model3
- …..
- model(n)

Second iteration – data, model estimation

- model1
- model2
- model3
- …..
- model(n)

Comparing modeling epochs:

1) can have same model with different parameters

OR

2) new model with different parameters.

Modeling **must check all models** and then via automation pick optimal model.

You need GPU server(s) to do this…..

Live Content Slide

When playing as a slideshow, this slide will display live content

Poll: When setting premium trend/reserve trend, I (we) currently incorporate our data science finding(s):

Live Content Slide

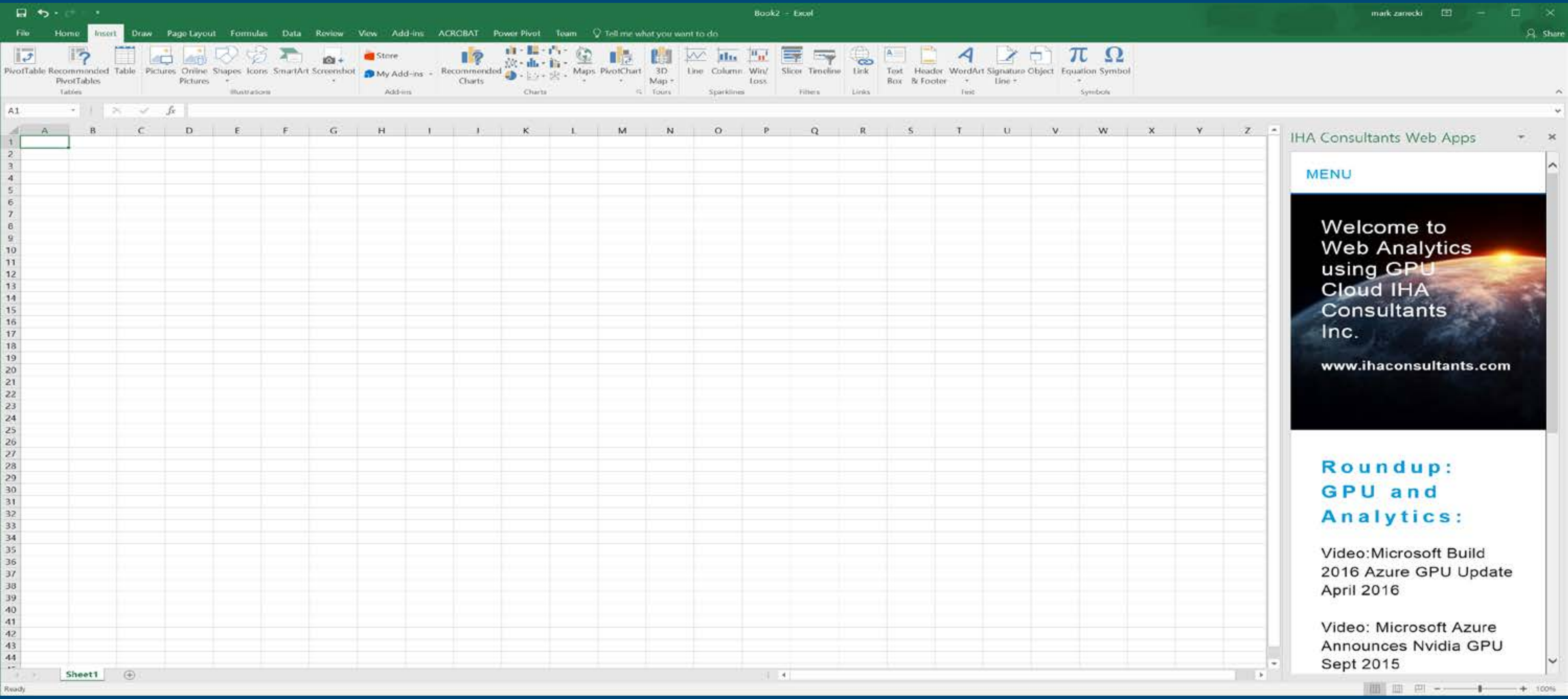
When playing as a slideshow, this slide will display live content

Poll: In the future, I expect data science to help me in my work by:

Demo : First Generation Predictive Loss Reserve Model

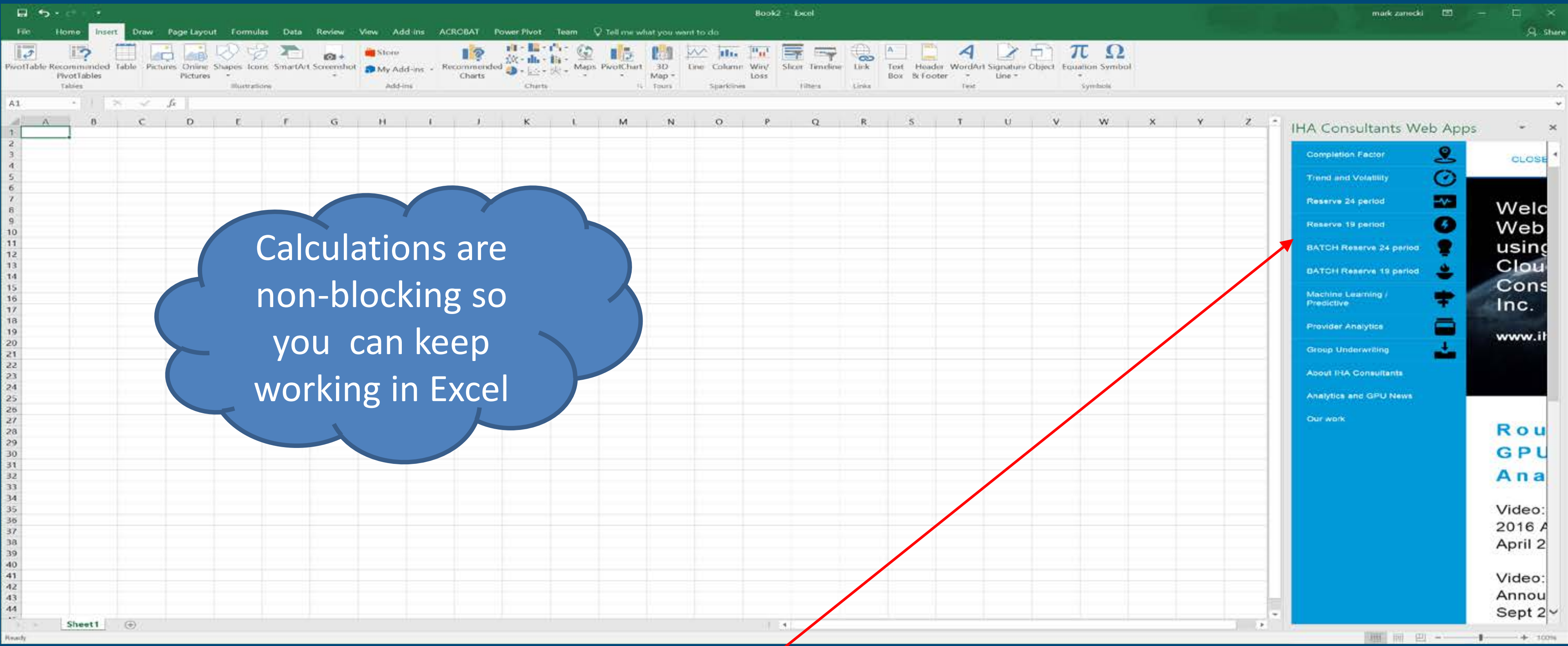
- We'll chose top-down predictive approach and model in (mean, variance) space using aggregate data. Can apply same modeling to predictive bin approach if desired.
- Reserve range is developed that is inclusive of manual techniques in compressed time.
- Free staff to perform detailed level analysis – "the why."
- "Hands-free" calculation. Can continue working in Excel on other worksheets or applications.
- The software offers seamless integration with existing systems with full data security.
- Custom software application which uses Excel as user interface (via web app functionality)with gpu server implementing predictive framework on backend. This is not a macro, not a DLL.
- User installable and configurable.
- Batch capability is a featured.
- Runs in cloud on gpu servers (no sharing of server vm or gpu card)
- Easy to use at any experience level.
- Can run multiple instances simultaneously for multi-tasking on same machine.

Demo : First Generation Predictive Loss Reserve Model



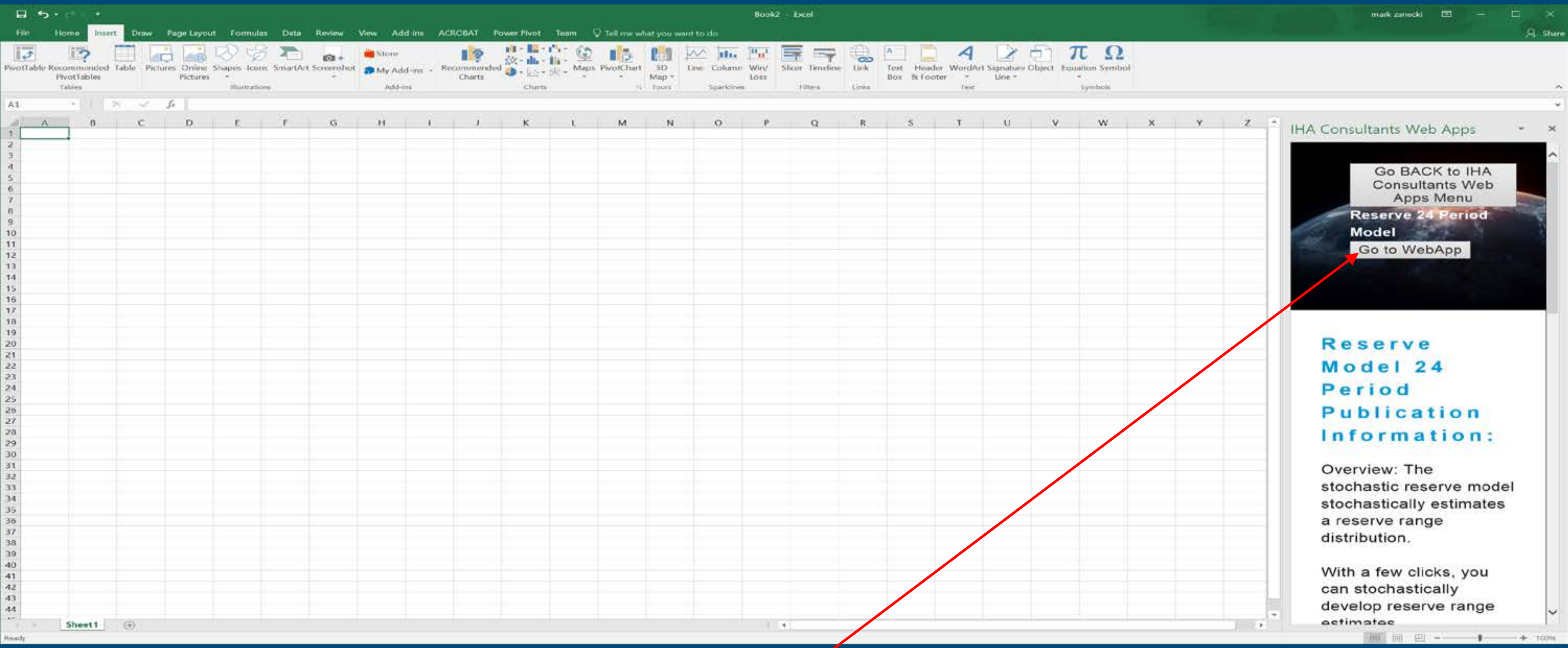
Welcome screen
Select 'Menu'

Demo : First Generation Predictive Loss Reserve Model



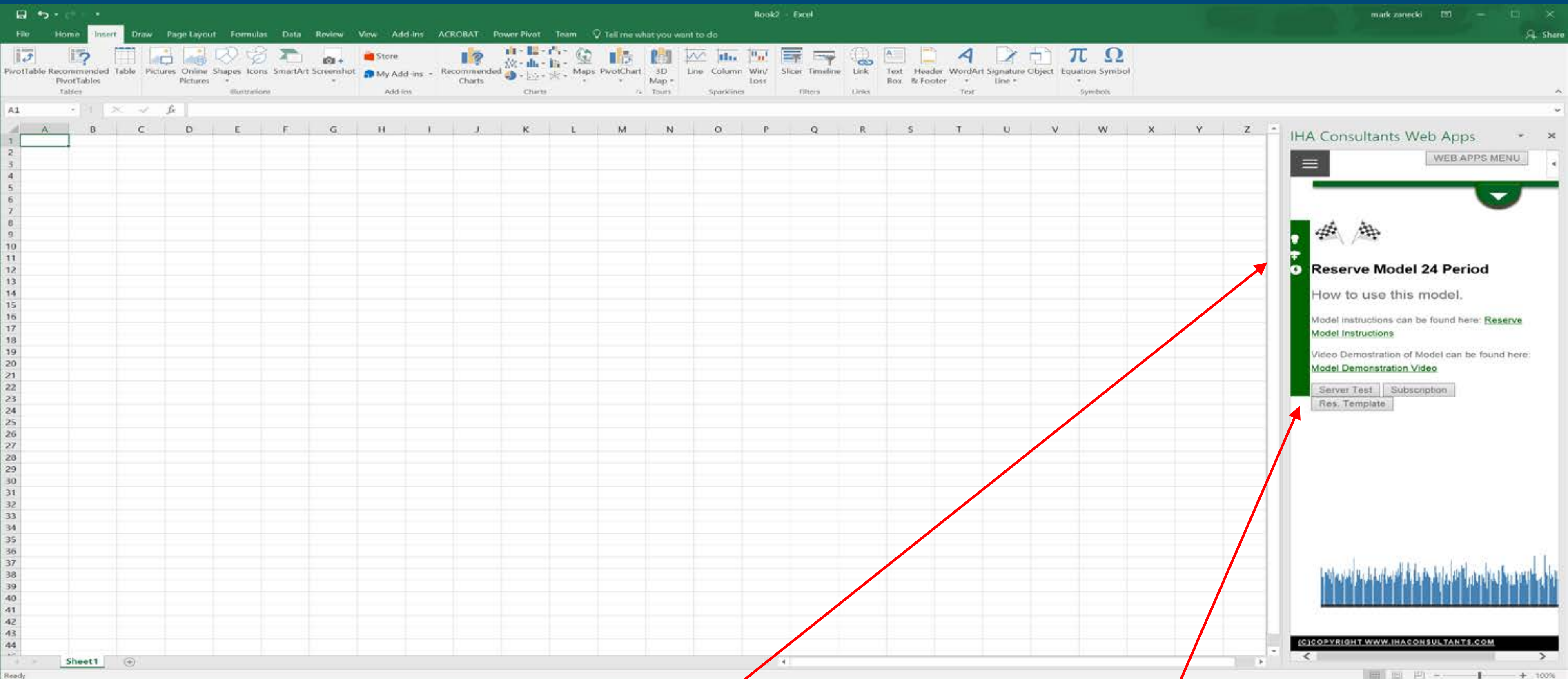
Click-on, ‘Select Reserve 24 period’ tab.
Functionality includes: Completion Factor, Trend & volatility, Reserve 19, Batch Reserve 24 & 19 and Machine Learning….

Demo : First Generation Predictive Loss Reserve Model



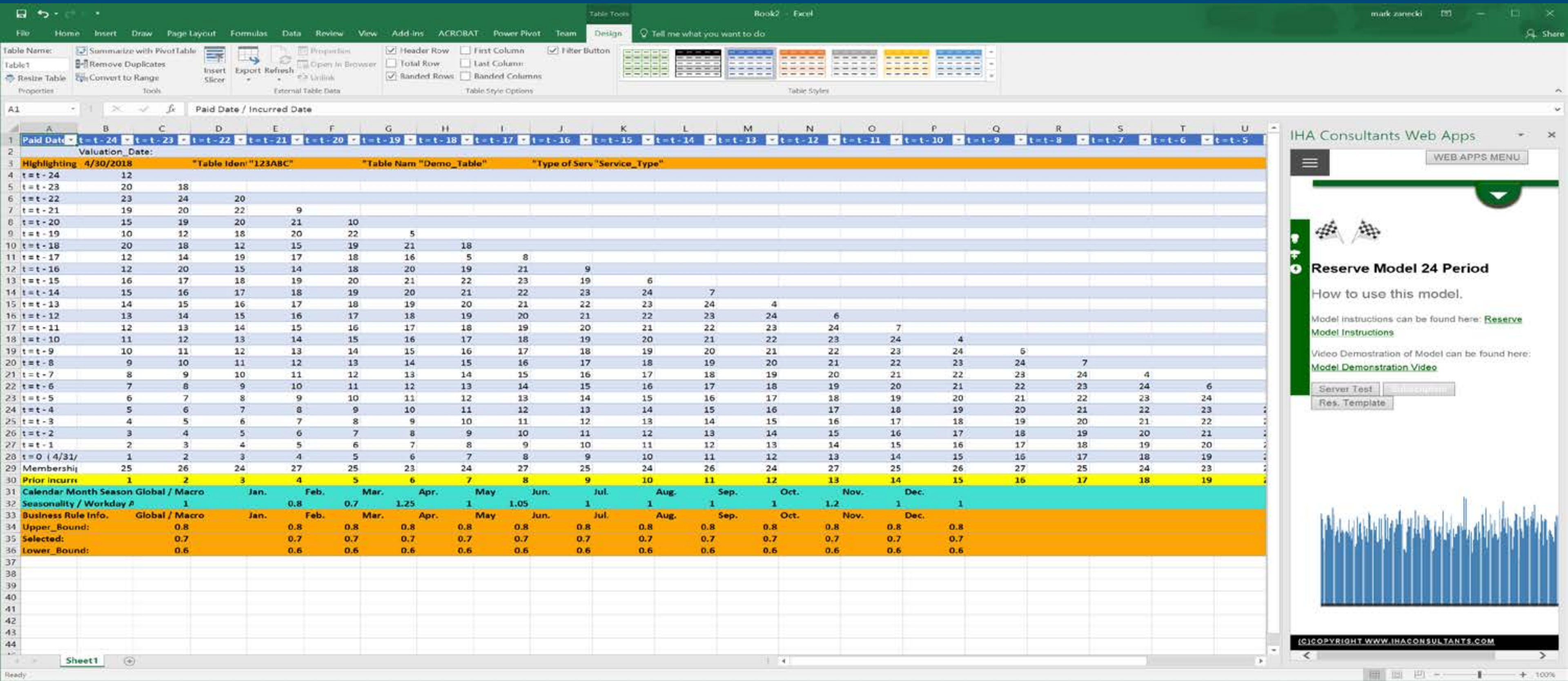
Click-on ‘Select Reserve 24 period’ tab.
Reveals Reserve 24 period product page with video and pdf instructions.
Click-on, ‘Go to WebApp’ button.

Demo : First Generation Predictive Loss Reserve Model



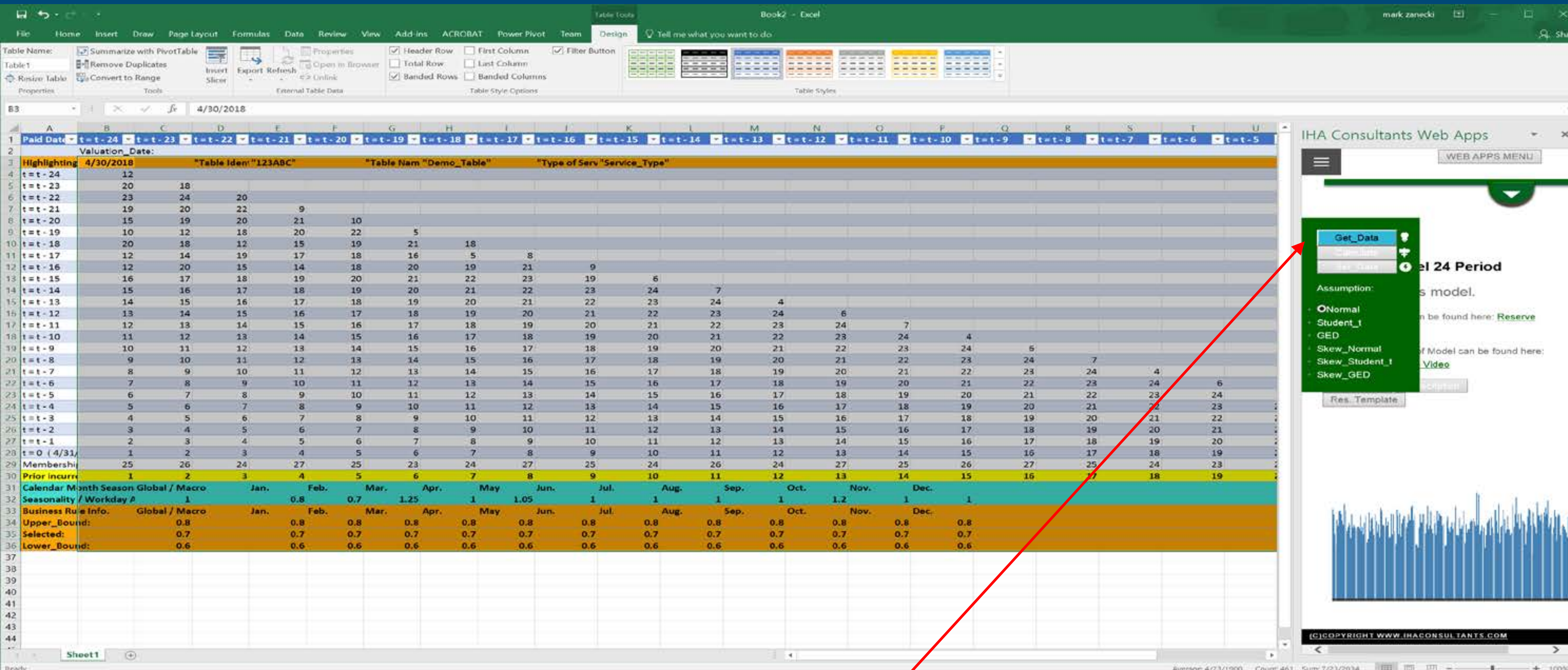
Reserve Model 24 period User Interface – interactive with Excel worksheet.
Orientation: Slide menu with icons containing action controls in green border.
Click-on, 'Res. Template' button to reveal data input template.

Demo : First Generation Predictive Loss Reserve Model



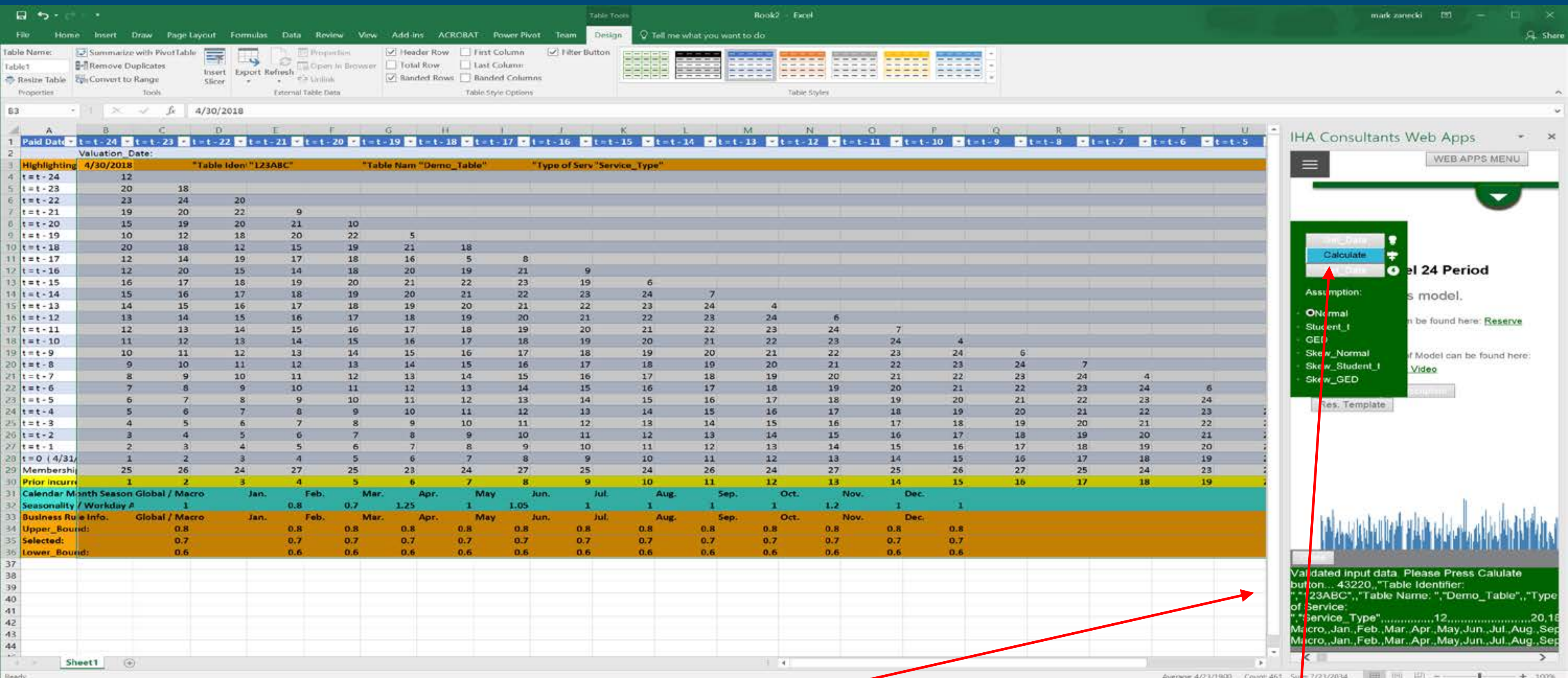
Copy and paste in data into template and then select range (B3:Z36)

Demo : First Generation Predictive Loss Reserve Model



With mouse select 'Get Data' on slide out menu.

Demo : First Generation Predictive Loss Reserve Model

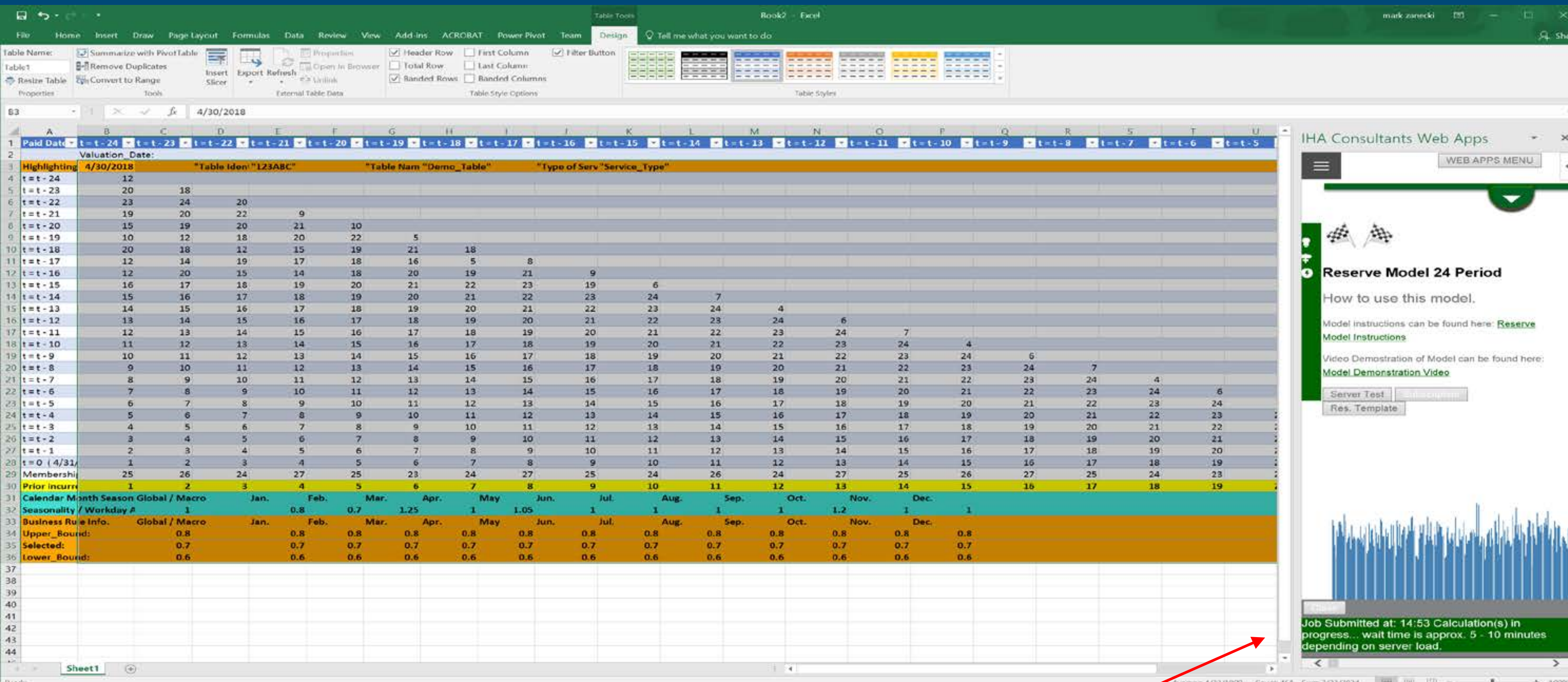


A data validation notification appears.

With mouse select ‘Calculate’ on slide out menu to perform value reserve table ‘hands-free.” A notification of calculation in progress appears.

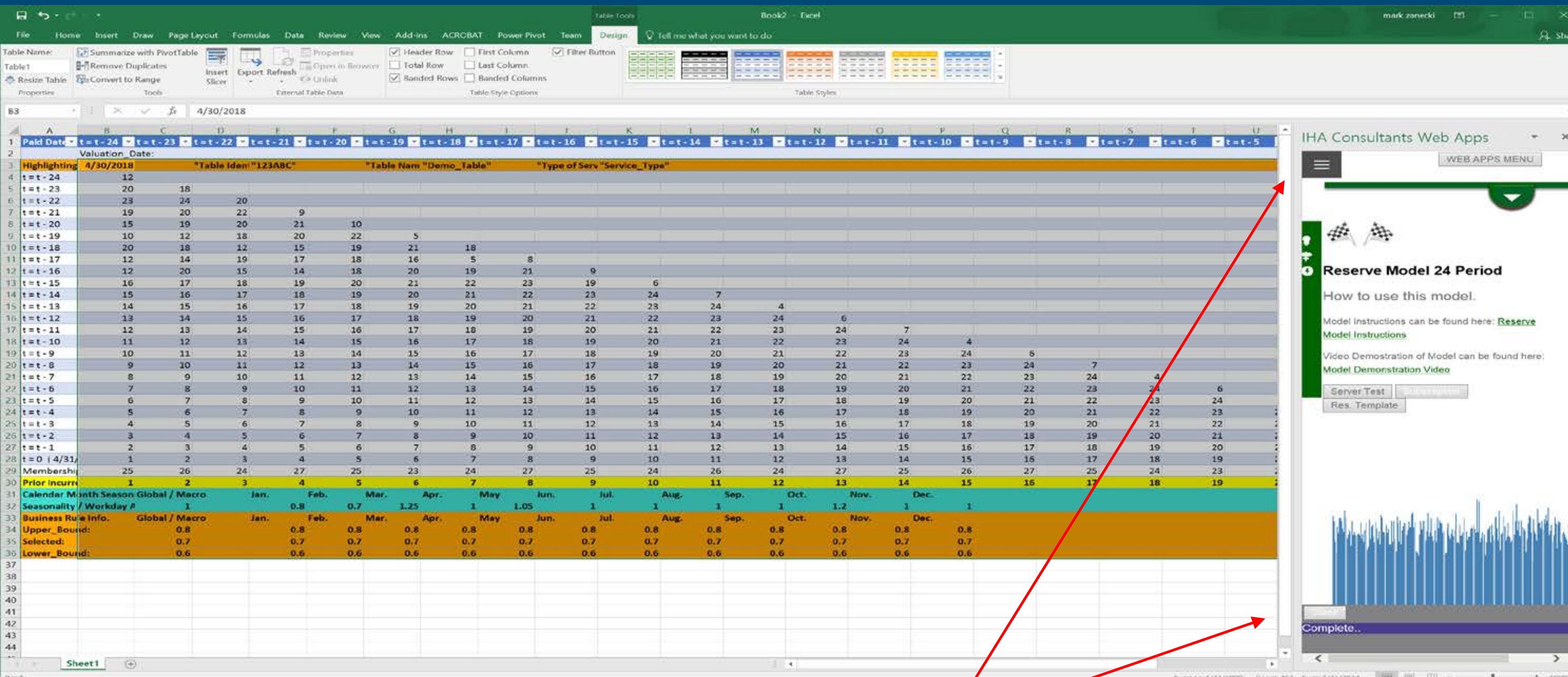
IHA Consultants Inc.

Demo : First Generation Predictive Loss Reserve Model



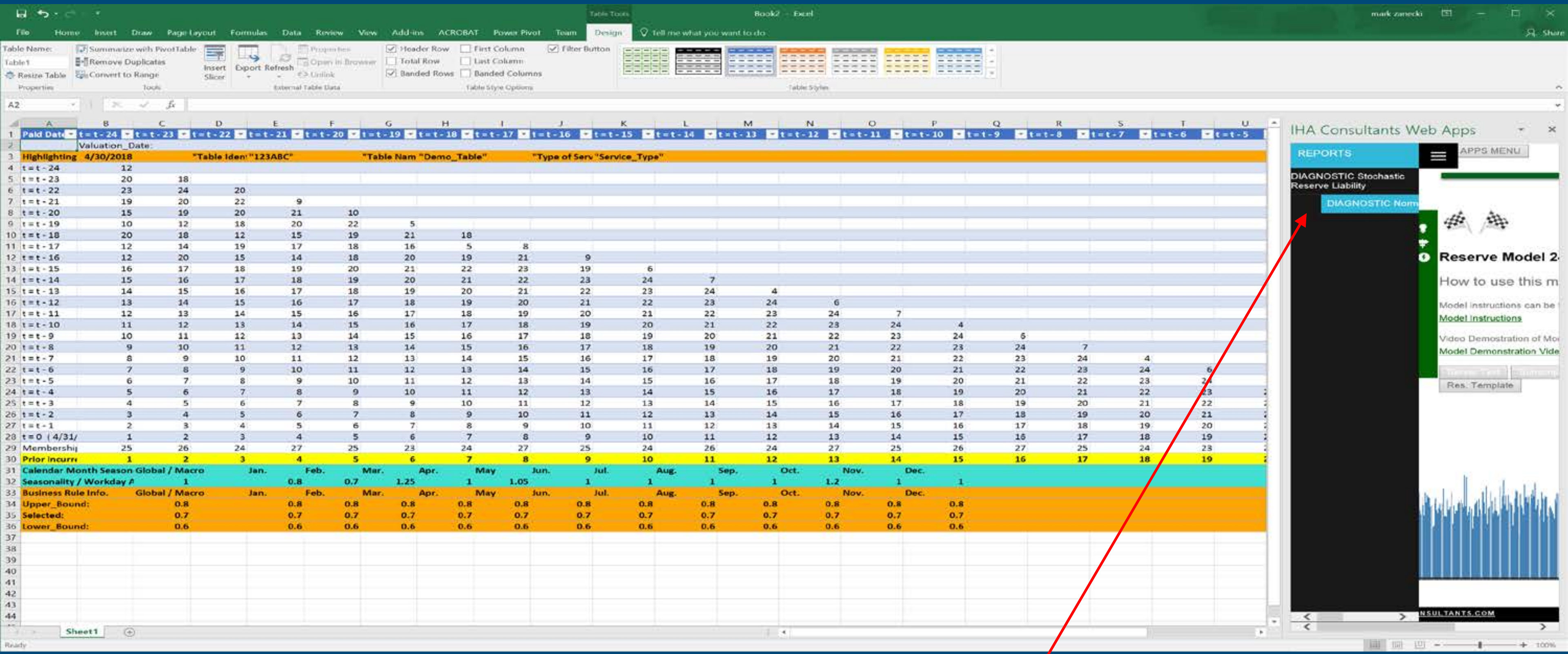
A notification of calculation in progress appears.
Can be 3 to 5 – 10 minutes depending on gpu virtual server.
Notification that results are ready will appear.

Demo : First Generation Predictive Loss Reserve Model



Notification that results are ready has appeared.
Click-on, ’hamburger icon’ to reveal available reports. (We will be exporting to worksheet.)

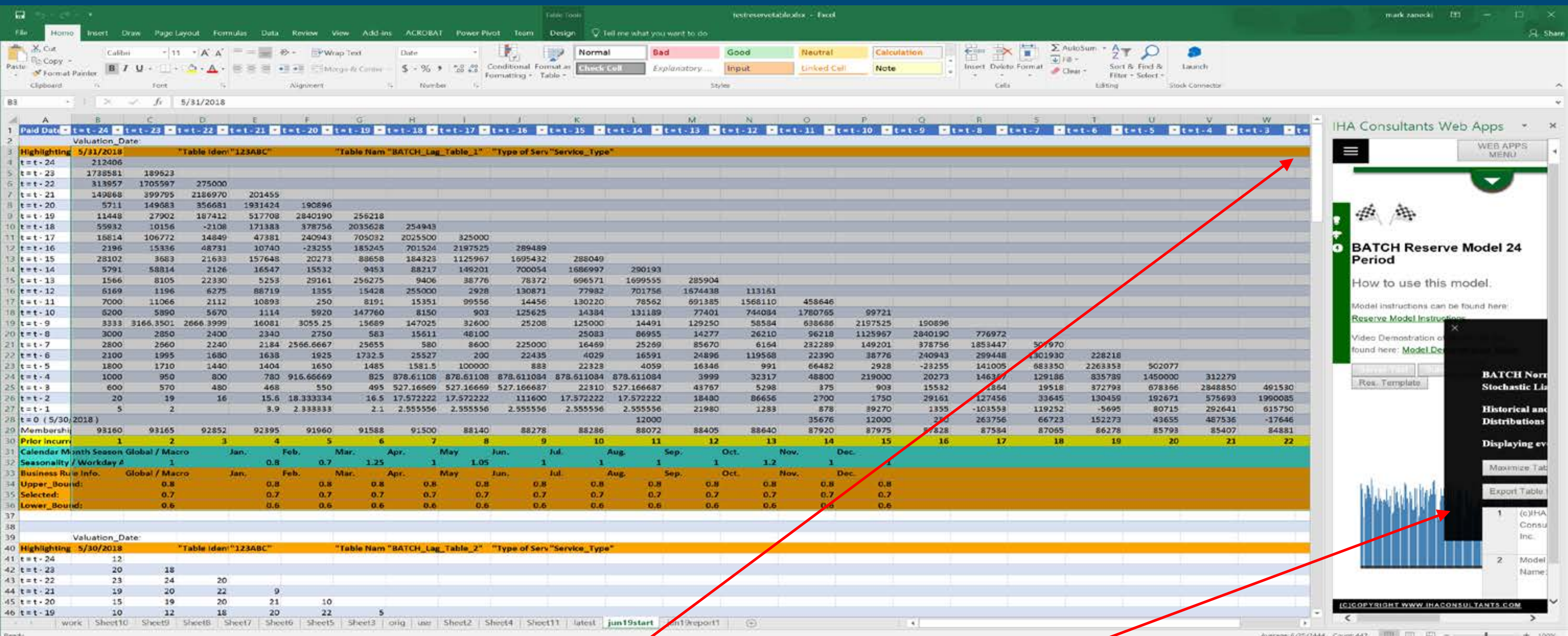
Demo : First Generation Predictive Loss Reserve Model



Click-on, ’hamburger icon’ to reveal available reports. (We will be exporting to worksheet.)
Click-on, ‘DIAGNOSTIC Normal’ button in slide out menu.

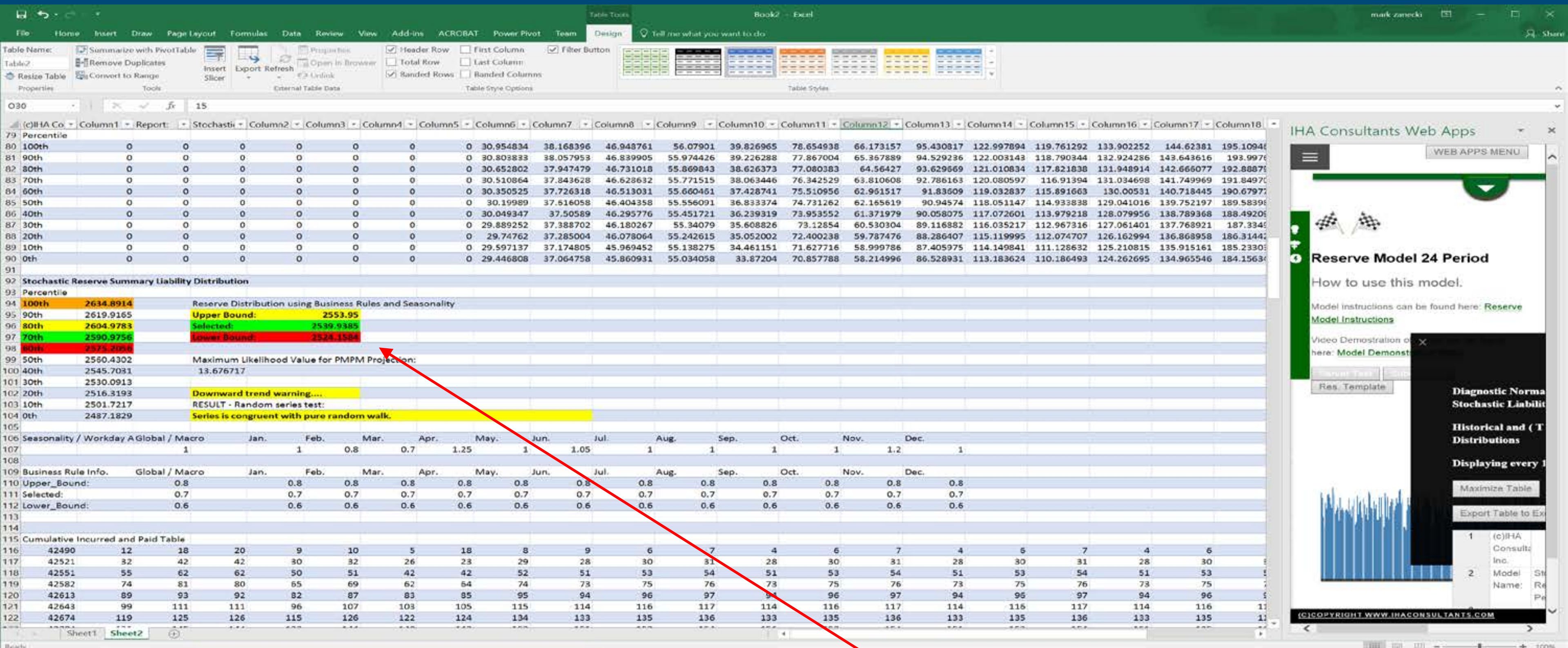
IHA Consultants Inc.

Demo : First Generation Predictive Loss Reserve Model



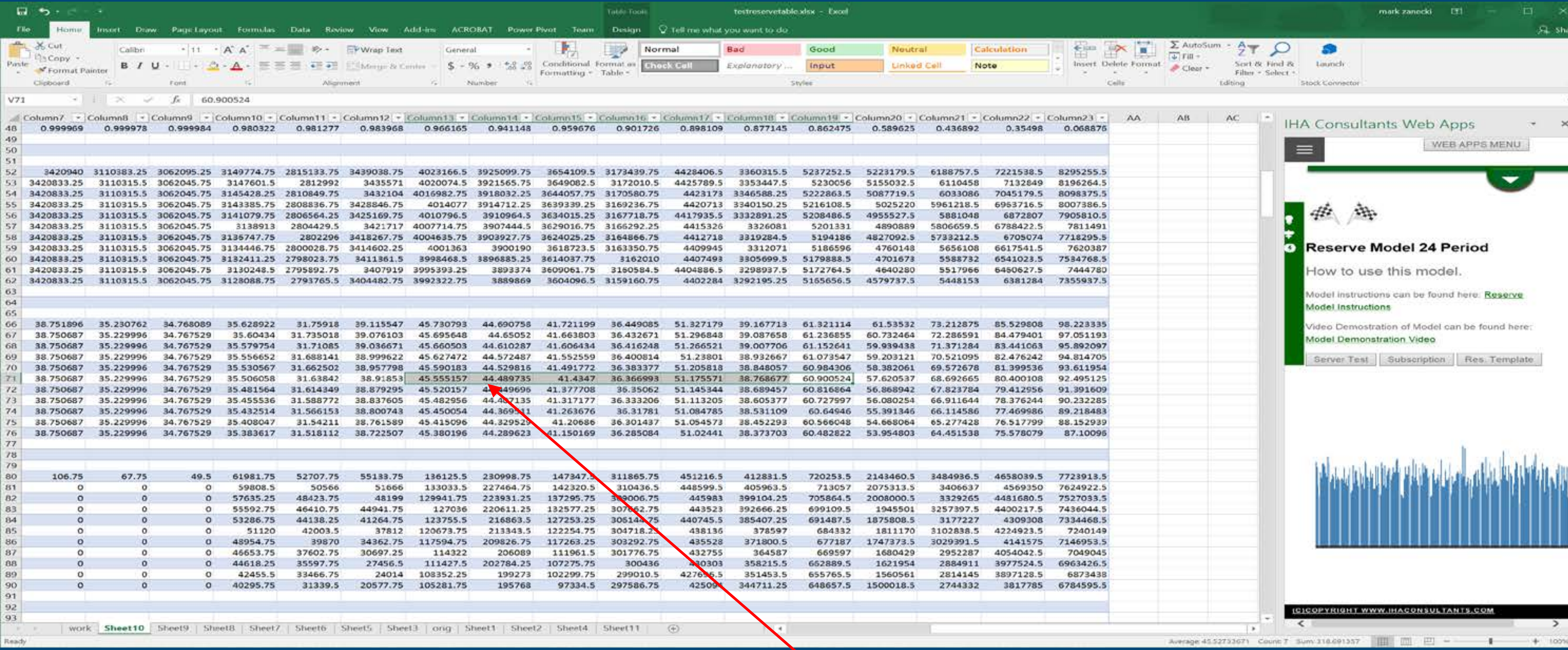
Click-on, ’hamburger icon’ to dismiss slide out menu.
You can now see the Reserve 24 valuation report.
Select a new tab. Next, click-on, ‘Export Table to Excel’ button.

Demo : First Generation Predictive Loss Reserve Model



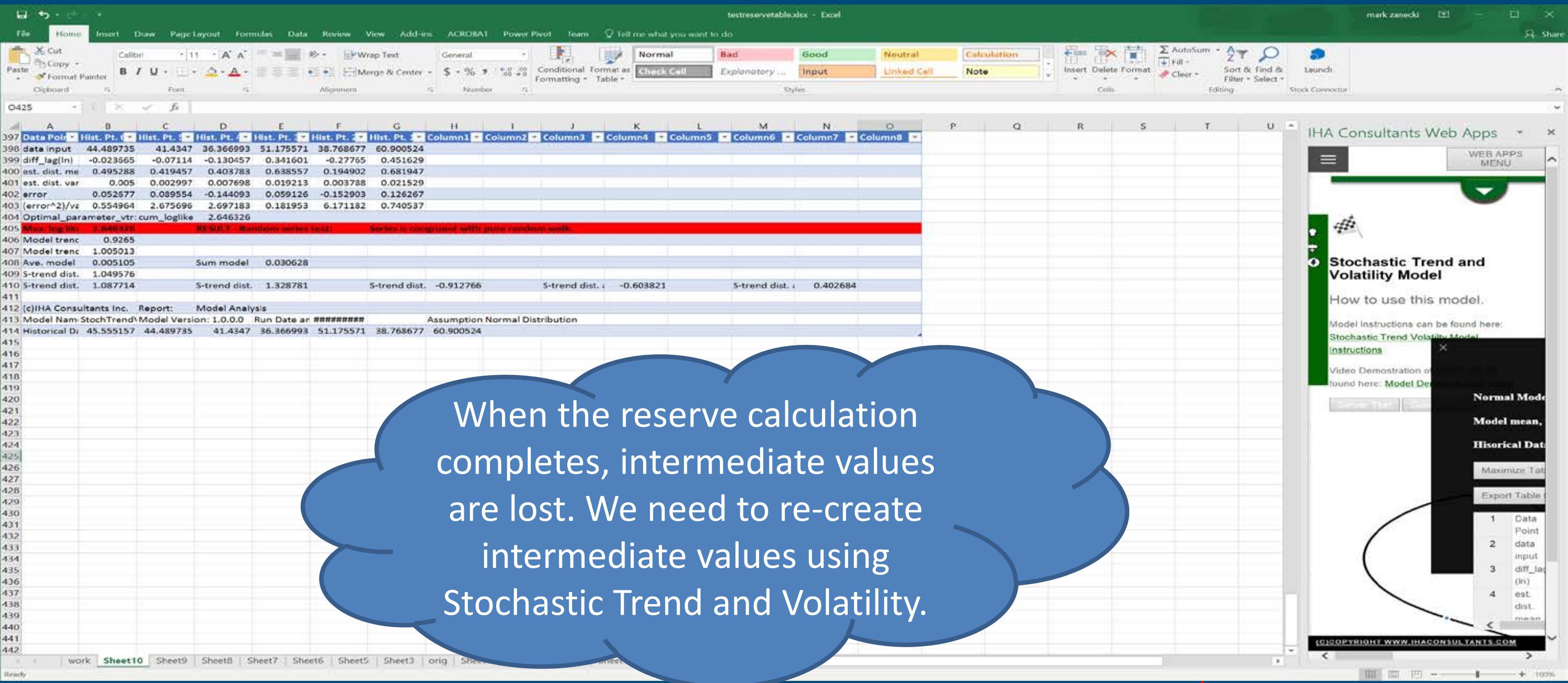
Scroll down until view Reserve 24 period range area in color highlights.
The gpu server has calculated thru first round of machine learning predictive models.
Diagnostic information is available below by scrolling down.

Demo : First Generation Predictive Loss Reserve Model



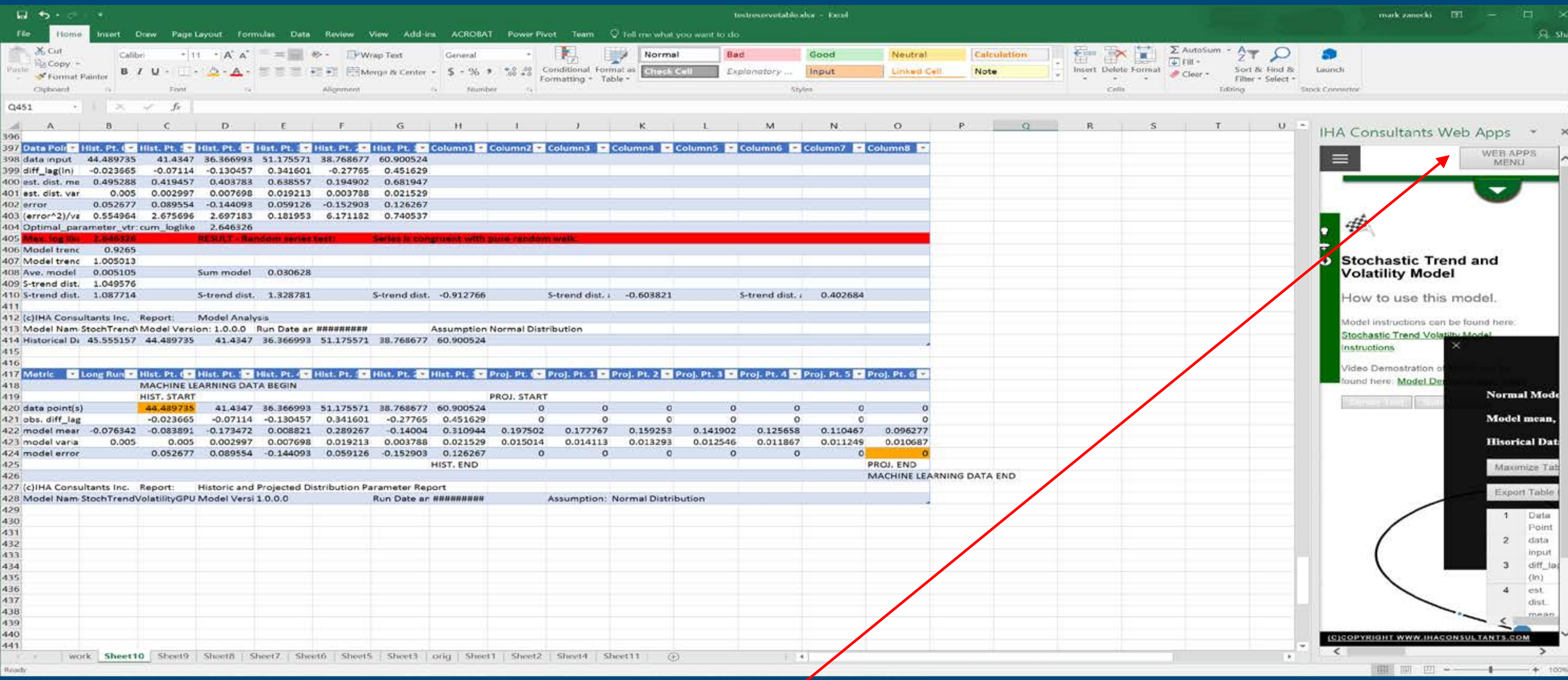
Scroll down to the last predictive calculation diagnostic area.
We’ll do a second round of machine learning on the projected pmpm durations.

Demo : First Generation Predictive Loss Reserve Model



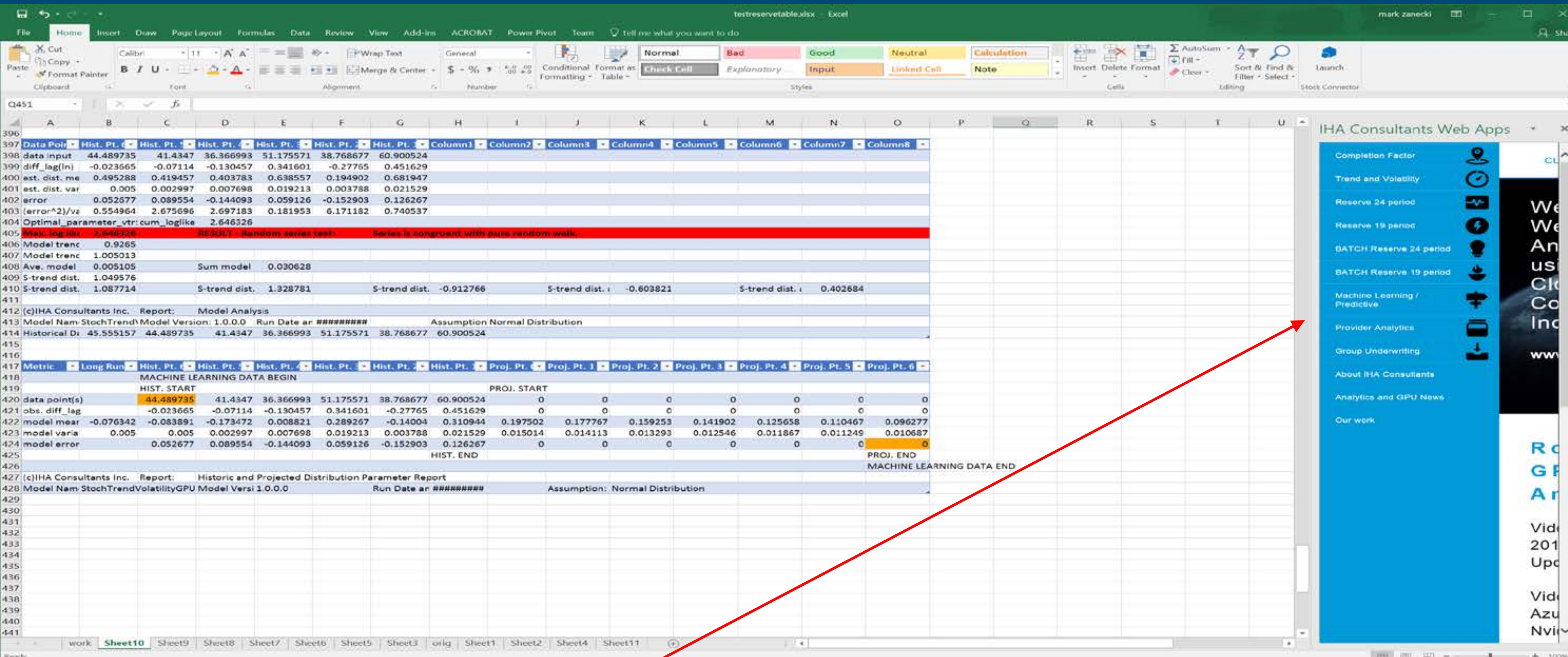
We'll need to go back to main menu and select 'Trend and Volatility' tab.

Demo : First Generation Predictive Loss Reserve Model



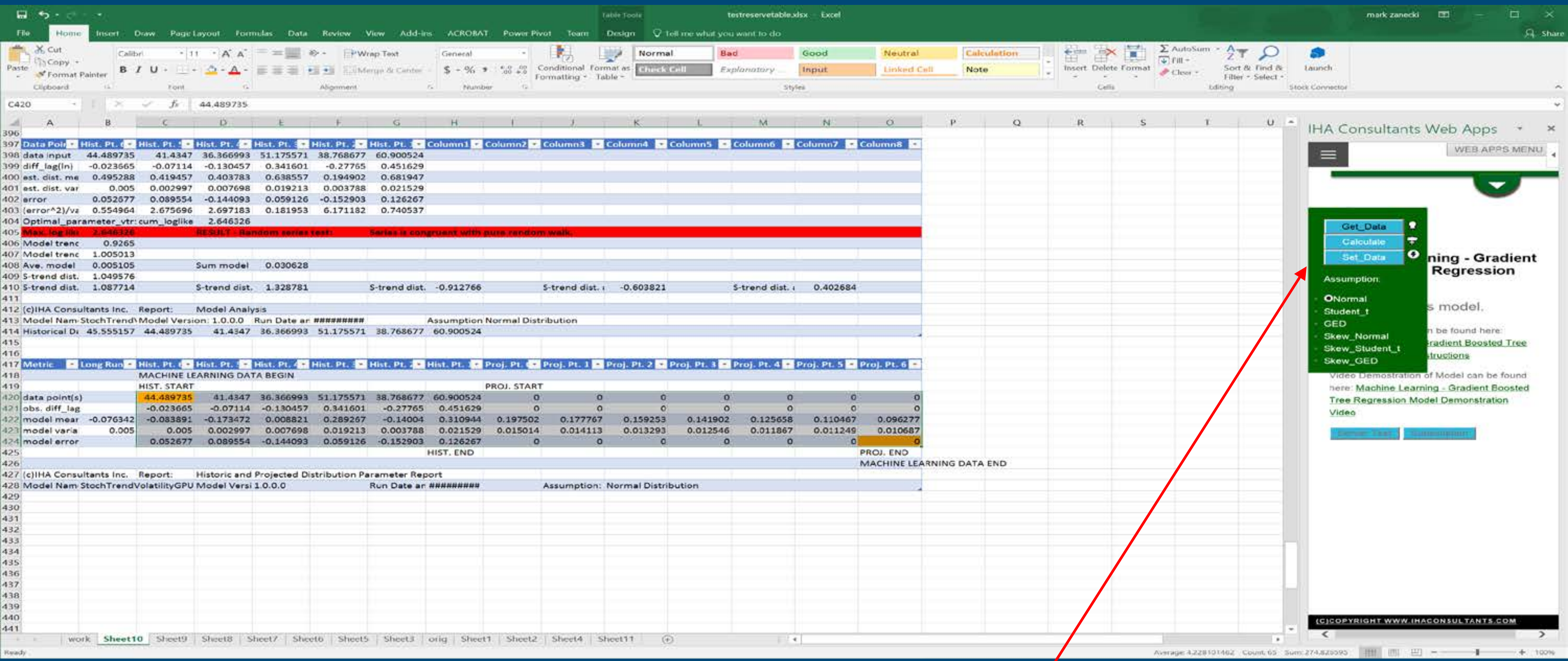
We exported the machine learning report and now can go back thru menu to ML tab. Click-on, 'Web-Apps Menu' button.

Demo : First Generation Predictive Loss Reserve Model



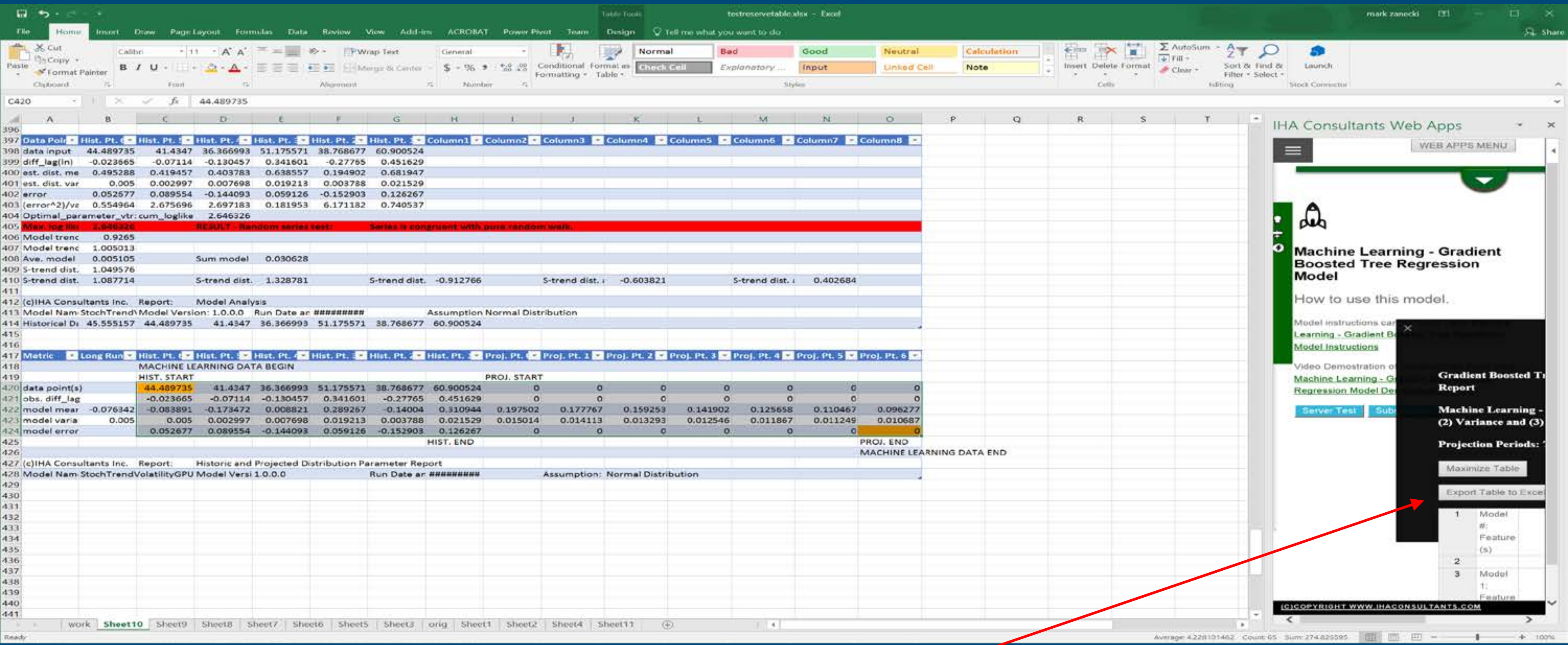
Click-on, ‘Machine Learning’ tab.

Demo : First Generation Predictive Loss Reserve Model



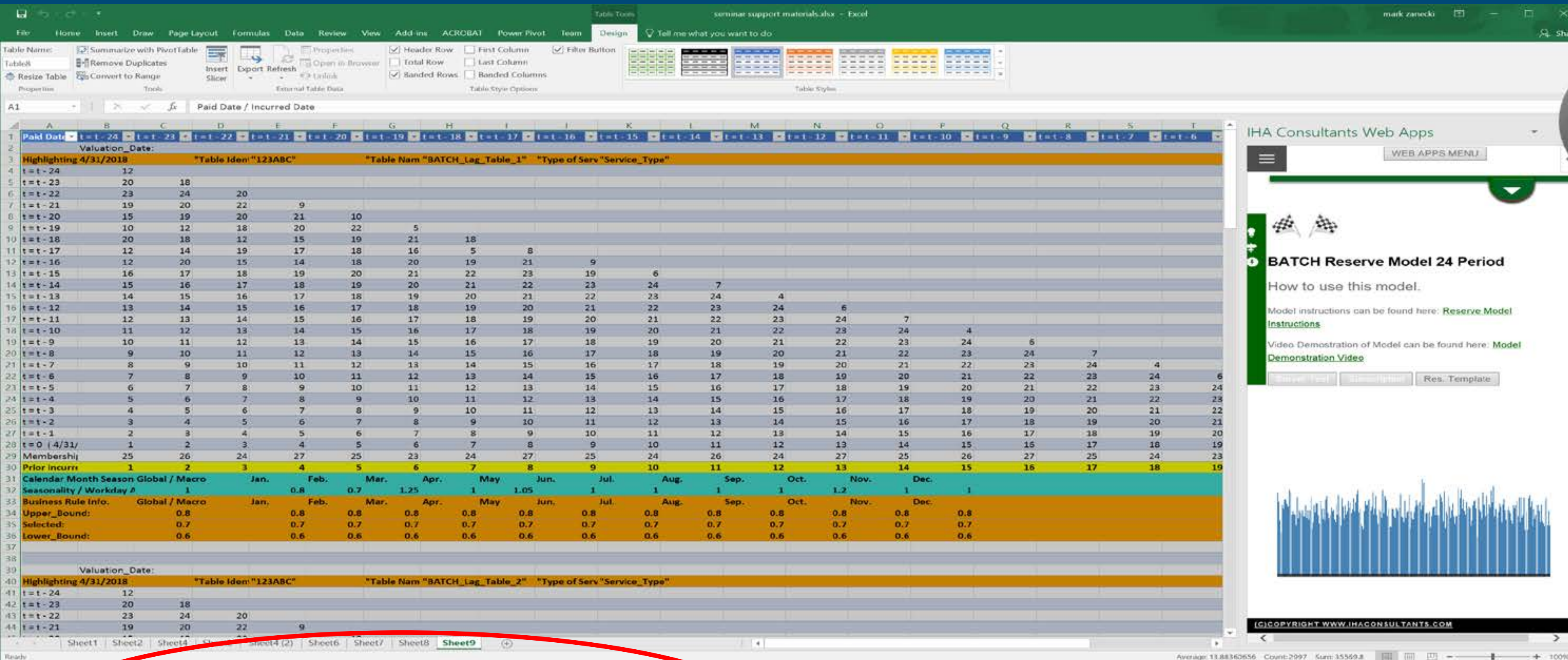
Highlight data. Select ‘Get Data’ and then press ‘Calculate.’

Demo : First Generation Predictive Loss Reserve Model



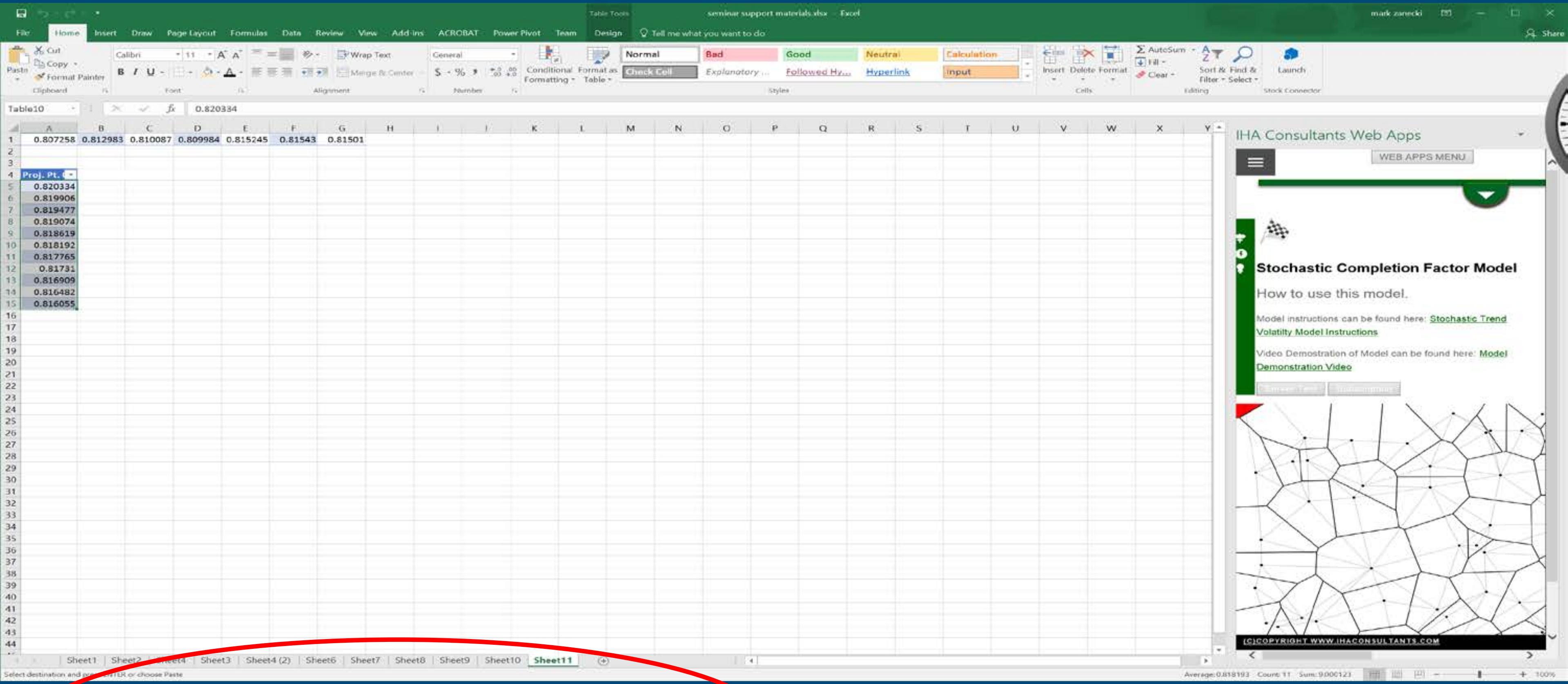
When calculation complete, click-on 'hamburger' and select machine learning value report. Export by pressing, 'Export to Excel.'
It was that easy and fast.

Demo : First Generation Predictive Loss Reserve Model



BATCH Reserve 24 with template shown.

Demo : First Generation Predictive Loss Reserve Model



Completion Factor

Applications of the Model

- These model(s) produce a distribution of reserve and ultimate liabilities ‘hands-free’ using layers of machine learning models.
- Batch processing is available. (Just use template(s))
- Rough range estimate can occur in any given time frame with number of sufficient gpu servers, data validated data - format ready for processing. (Day 2,3)
- While reserve table calculations are “in flight,” detailed data analysis can occur looking for anomalies up front rather than reacting at end of development process.
- Seem-less integration with Excel allow productivity day 1.
- Data is securely encrypted at all times.
- The system is user installable in 5 – 10 minutes.
- Can use as stand alone or as supplement.
- Integration with pricing provided via Stochastic Trend and Volatility web-app and as well via Machine Learning web-app.

Summary

- Presented overview of current loss reserve modeling techniques
- Discussed predictive modeling approaches at model level and at data level.
- Equivalence of three levels if use proper predictive modeling technique for the particular level.
- Provided demo of top-down predictive modeling technique using (mean, variance) on aggregate data producing estimate range.
- Framework supports predictive modeling for loss reserve and pricing applications.
- Framework is non disruptive to current processes or IT infrastructure.

Any Final Questions?

Mark Zanecki, ASA, MAAA

Tel: 919 260-3291

mark.zanecki@ihaconsultants.com

www.ihaconsultants.com