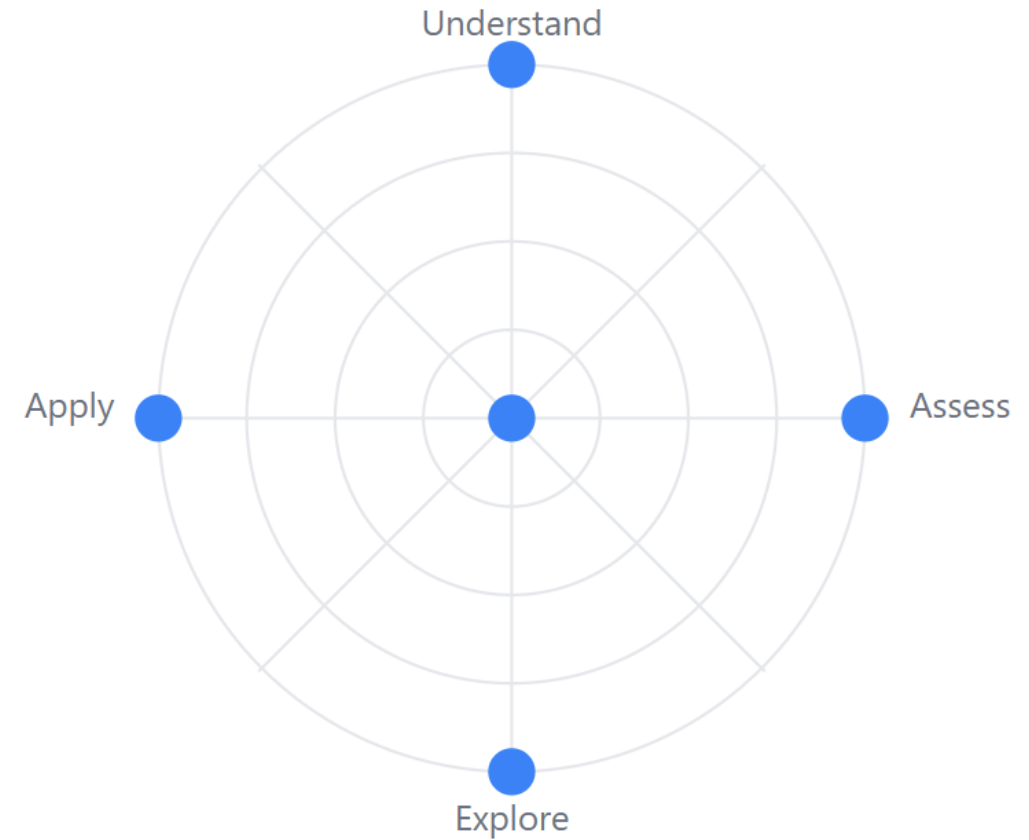

Independent Validation

Qualitative Aspects

Workshop Session 1

Workshop Objectives

- ✓ Understand the qualitative subcomponents of independent validation
- ✓ Learn how to assess these subcomponents using practical methods
- ✓ Explore real-world examples
- ✓ Apply these concepts to your teams's models



Workshop Structure

01

Conceptual Soundness

Evaluating theoretical foundations and design principles

02

Performance Monitoring

Ensuring sustained accuracy and relevance over time

03

Model Testing

Examining qualitative aspects of test design and execution

04

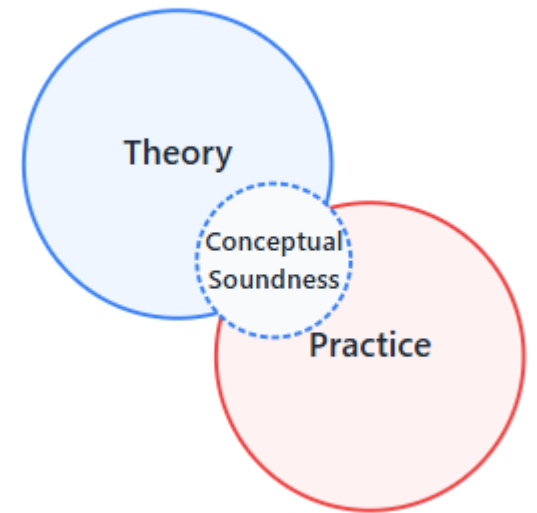
Governance & Controls

Managing review processes and organizational oversight

Conceptual Soundness and Design

Overview

Conceptual Soundness and Design evaluates whether a model's theoretical foundation is appropriate for its intended use. This qualitative process involves assessing the model's theory, development process, assumptions, limitations, and alternative considerations.



Conceptual Soundness and Design

Model Theory and Approach

Assess whether the mathematical, economic, or business theory underpinning the model is appropriate for the intended use case.

Key Assumptions

Evaluate the reasonableness and impact of assumptions made during model development, including their implications for model outcomes.

Limitations and Weaknesses

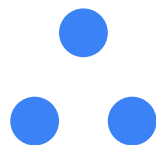
Identify and assess potential limitations, edge cases, and scenarios where the model may not perform as expected.

Alternative Approaches

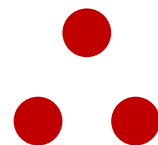
Consider other modeling techniques that could address the same business problem and evaluate why the current approach was selected.

Model Theory and Approach

A

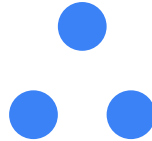


B

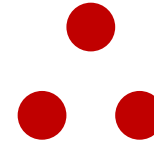


Model Theory and Approach

A



B



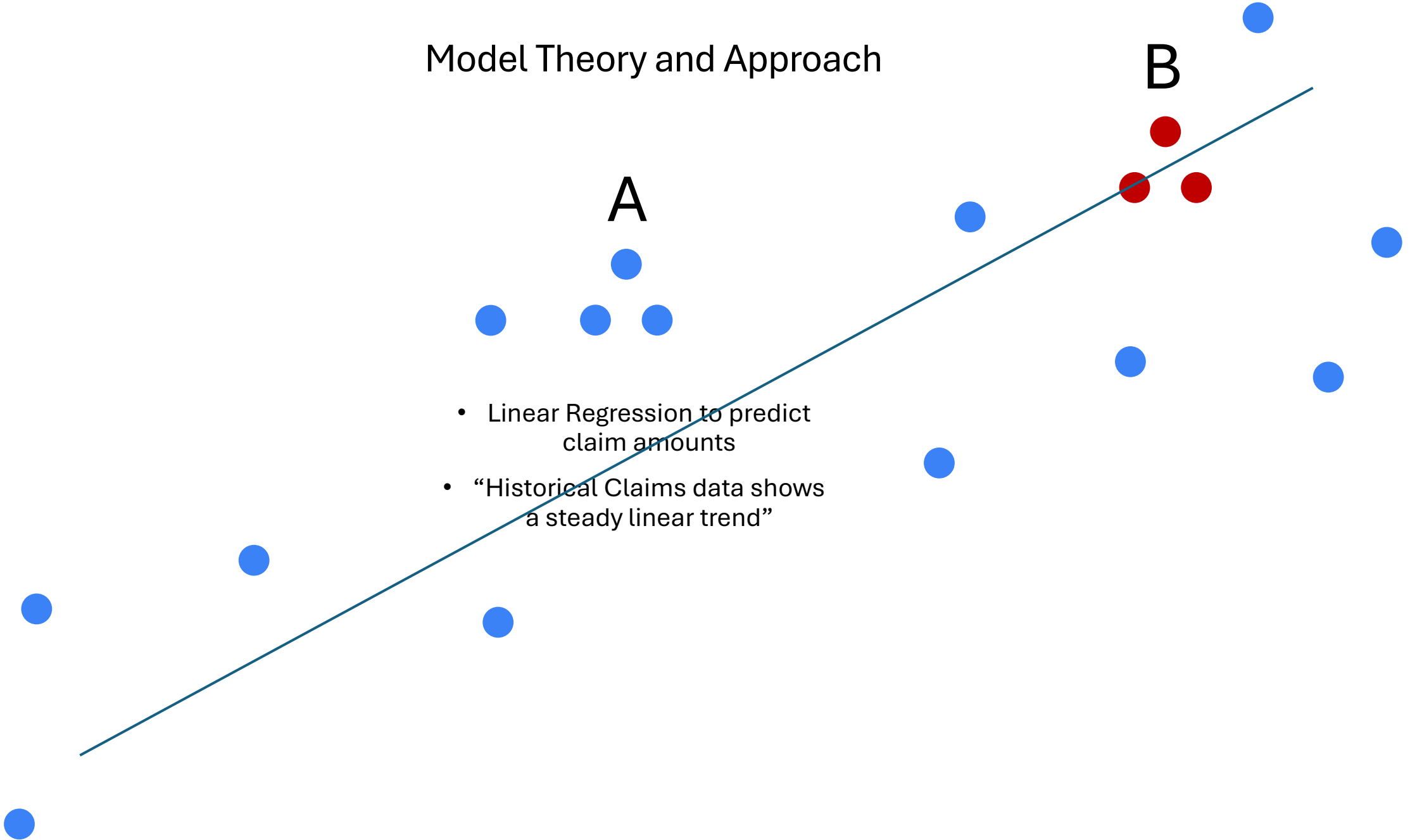
- Linear Regression to predict claim amounts
- “Historical Claims data shows a steady linear trend”

Model Theory and Approach

A

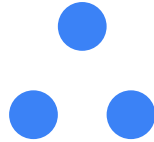
B

- Linear Regression to predict claim amounts
- “Historical Claims data shows a steady linear trend”

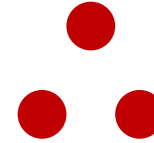


Model Theory and Approach

A



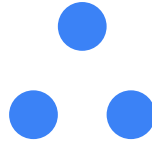
B



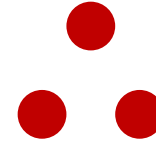
- Linear Regression to predict claim amounts
- “Historical Claims data shows a steady linear trend”

Model Theory and Approach

A



B



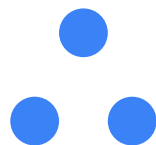
- Linear Regression to predict claim amounts
- “Historical Claims data shows a steady linear trend”



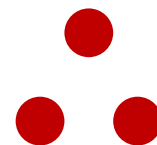
The theory matches the problem—linear regression suits their data’s pattern—and they justify it with evidence.

Model Theory and Approach

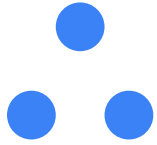
A



B

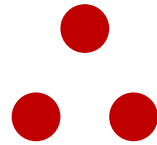


A



Model Theory and Approach

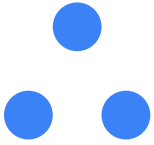
B



- They use a random forest for churn prediction
 - They do not give an explanation

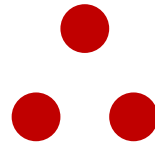
Churn prediction is the process of estimating which customers are likely to stop using a product or service

A



Model Theory and Approach

B



- They use a random forest for churn prediction
 - They do not give an explanation



Without explaining why random forest fits churn prediction (capturing non-linear patterns), validators can't assess its appropriateness

Churn prediction is the process of estimating which customers are likely to stop using a product or service

Key Assumptions

- Data trends and historical patterns

- *Example:* Assume a 0.05 significance level for regression coefficients, with sensitivity analysis to confirm robustness.

- Model Parameters

- *Example:* Assume inflation impacts repair costs, supported by economic forecasts from credible sources.

- External Factors and Market Conditions

- *Example:* Assume customer age is a key churn driver, corroborated by feature importance in a random forest model.

- Customer Behavior and Demographics

- *Example:* Assume customer age is a key churn driver, corroborated by feature importance in a random forest model.

- Vendor or Third-Party Selection

- *Example:* Assume a vendor's actuarial software is reliable, based on their track record and contract terms.

Limitations and Weaknesses

- Rare or Extreme Events

- *Example: Linear regression may struggle with high-cost claims (e.g., totaled luxury vehicles); recommend quarterly outlier monitoring.*

- Data Volume and Quality

- *Example: Random forest may overfit with a small dataset (e.g., 1,000 customers); suggest periodic retraining with larger samples.*

- Data Relevance and Timeliness

- *Example: Model performance may degrade if claims data excludes recent electric vehicle trends.*

- Model Complexity and Fit

- *Example: Linear regression may miss non-linear claim spikes; consider complementary models for validation.*

- Vendor or System Dependencies

- *Example: Reliance on a vendor's outdated algorithm could skew predictions; monitor vendor updates regularly.*

Alternative Approaches

1. Could a less complex model suffice?
2. Could a more sophisticated method improve performance?
3. Are there other techniques better suited to your data?

Statistical -> Machine Learning
Machine Learning -> Statistical

Conceptual Soundness and Design

Document Checklist



- ☒ _____
- ☒ _____
- ☒ _____
- ☐ _____

- ☒ **Theoretical foundation** - Documentation of the mathematical theory and statistical concepts
- ☒ **Assumptions catalog** - Explicit listing of all assumptions with justifications
- ☒ **Limitations description** - Clear articulation of model constraints and edge cases
- ☐ **Alternative Approaches** - Documentation of the selection and evaluation of the current approach

Ongoing Performance Monitoring

Qualitative Subcomponents

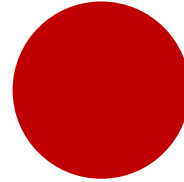
- 1 Data Quality Assessment**
Evaluate processes ensuring data completeness and accuracy
- 2 Data Appropriateness**
Assess data suitability for the model's purpose
- 3 Model Reporting**
Examine reports and communications about model performance

Data Quality



Focuses on the integrity of the data itself—ensuring it is complete (no missing values) and accurate (no errors)

Data Appropriateness



Evaluates whether the data fits the model's specific purpose and current context

Data Quality



Are data quality processes and updates clearly defined?

How does the team handle Null Values?

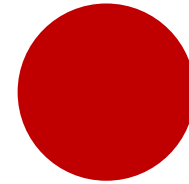
Were transformations done to data?

Data Appropriateness

Is there evidence of data relevance checks and regular reviews?

Does the team use outdated data?

Documentation of vendor selection



Evaluates whether the data fits the model's specific purpose and current context

Model Testing

Qualitative Subcomponents



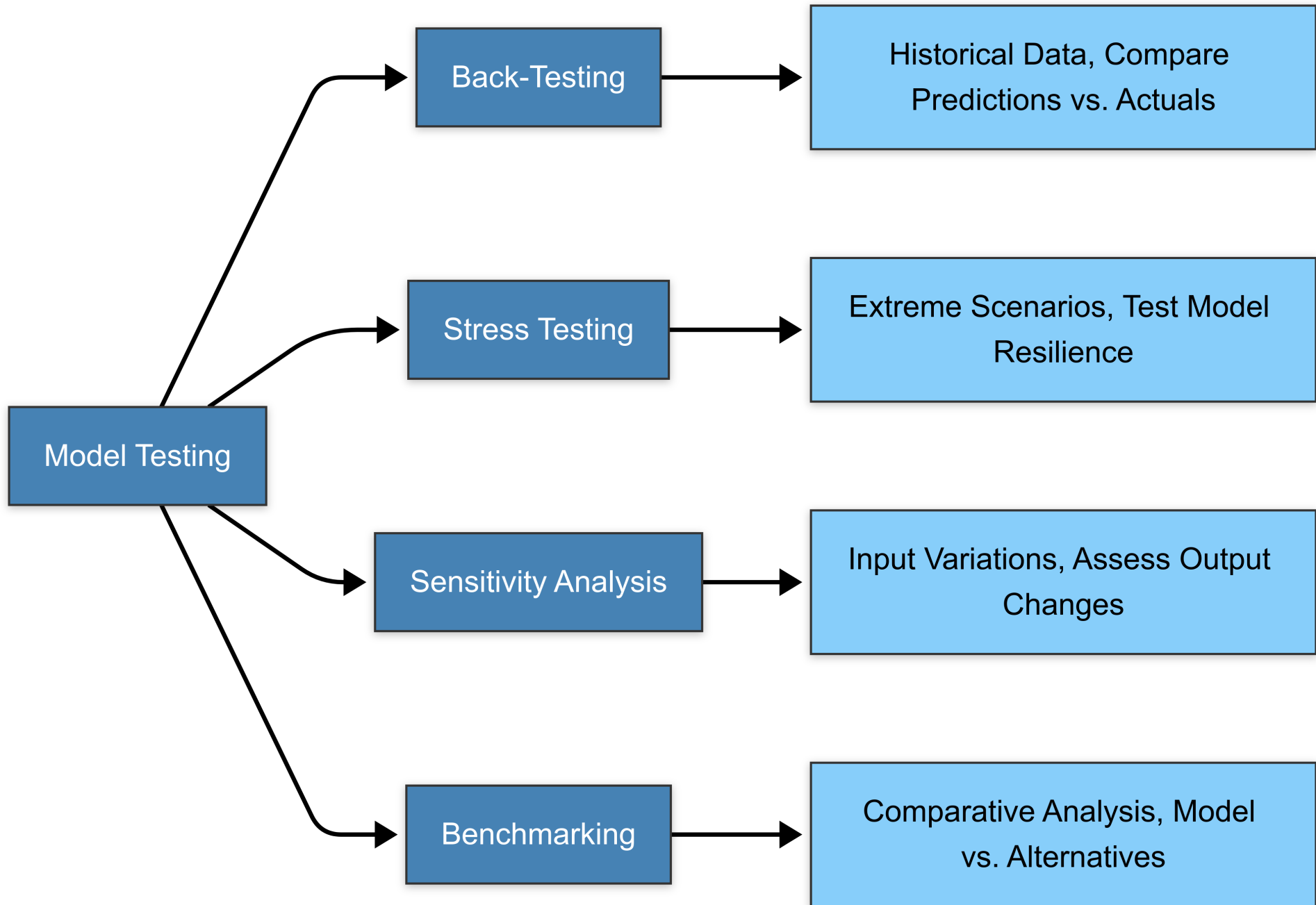
Test Design and Rationale

Evaluate why specific tests were chosen and their alignment with the model's goals and risk profile



Documentation of Testing

Review the clarity, completeness, and traceability of testing records and their accessibility to stakeholders





```
graph LR; A[Back-Testing] --> B[Historical Data, Compare Predictions vs. Actuals]
```

Back-Testing

Historical Data, Compare
Predictions vs. Actuals

- *Test Design and Rationale:* Why this test? Does it align it with your goal?
- *Example:* Back-test a regression model with 2023 claims to ensure pricing aligns with real costs.



```
graph LR; C[Stress Testing] --> D[Extreme Scenarios, Test Model Resilience]
```

Stress Testing

Extreme Scenarios, Test Model
Resilience


```
graph LR; A[Stress Testing] --> B[Extreme Scenarios, Test Model Resilience]; C[Sensitivity Analysis] --> D[Input Variations, Assess Output Changes];
```

Stress Testing

Extreme Scenarios, Test Model
Resilience

- *Test Design and Rationale:* Why this test? Match it to your risk profile.
- *Example:* Stress-test a random forest with a 20% premium hike to evaluate churn resilience.

Sensitivity Analysis

Input Variations, Assess Output
Changes

Sensitivity Analysis



```
graph LR; A[Sensitivity Analysis] --> B[Input Variations, Assess Output Changes]; C[Benchmarking] --> D[Comparative Analysis, Model vs. Alternatives];
```

Input Variations, Assess Output Changes

- *Test Design and Rationale:* Why this test? Target key assumptions.
- *Example:* Vary regression input (e.g., repair cost inflation) by $\pm 2\%$ to test pricing stability.

Benchmarking

Comparative Analysis, Model vs. Alternatives

```
graph LR; A[ ] --> B[Benchmarking]; B --> C[Comparative Analysis, Model vs. Alternatives];
```

Benchmarking

Comparative Analysis, Model
vs. Alternatives

- *Test Design and Rationale:* Why this test? Ensure competitiveness.
- *Example:* Benchmark a random forest churn model against a logistic regression baseline.

Documentation of Testing

1. Can an outsider understand the documentation?
2. Does the documentation rely solely on visuals, or does it balance visuals with written explanations
3. Are all critical details included, such as test methodology, data sources, assumptions, results, and stakeholder feedback

Governance and Controls

Qualitative Subcomponents

Policies and Procedures

Assess the comprehensiveness of model-related policies and their implementation

Model Documentation

Evaluate the quality, completeness, and accessibility of model documentation

Change Management

Review processes for updating models and tracking model versions

Model Administration

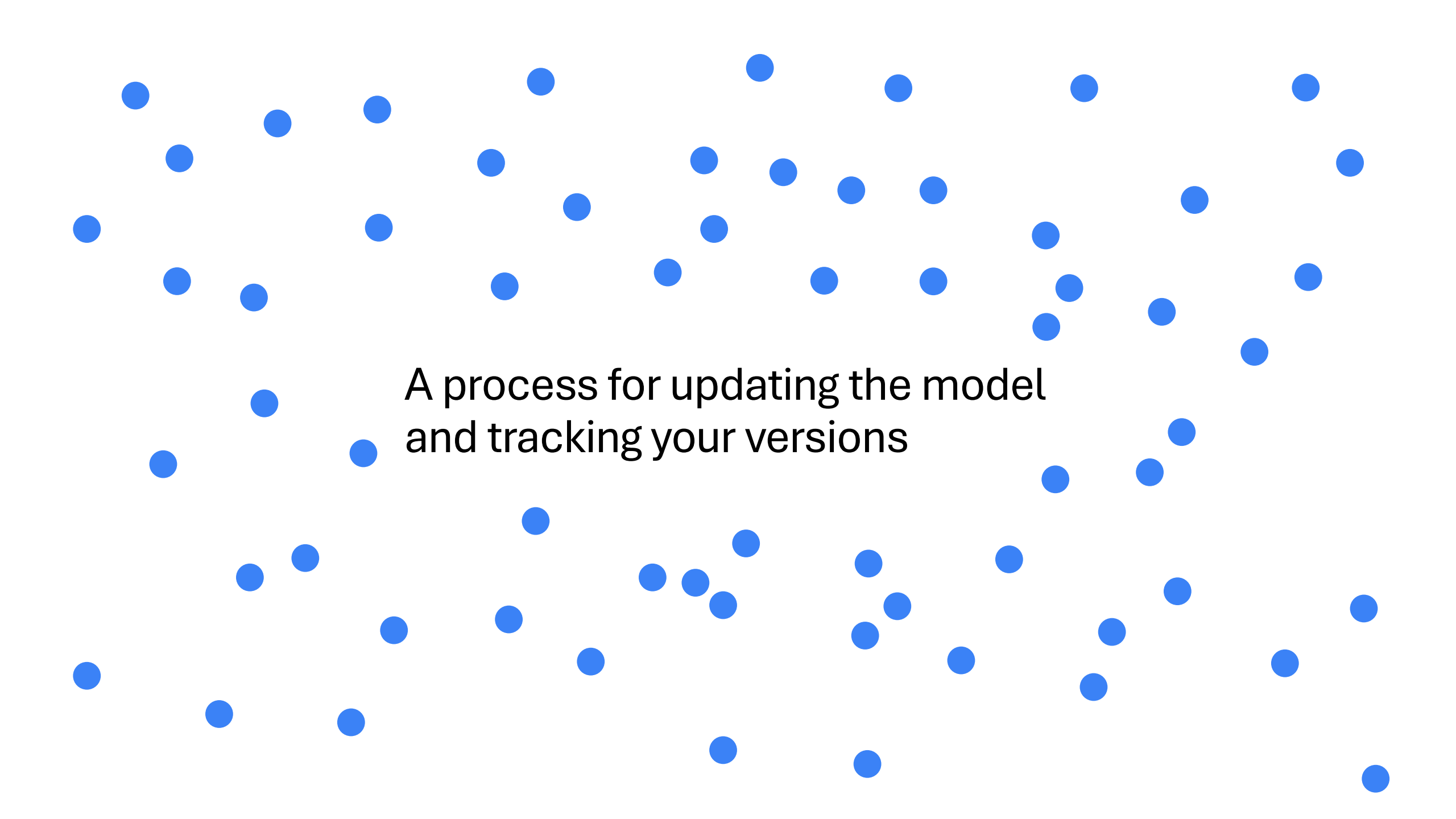
Assess controls on access, usage, and roles/responsibilities

The background of the slide is white and filled with numerous small, solid blue circles scattered in a random pattern. These dots are of uniform size and color, creating a decorative, starry effect around the central text.

Do you have a Change
Management Procedure?

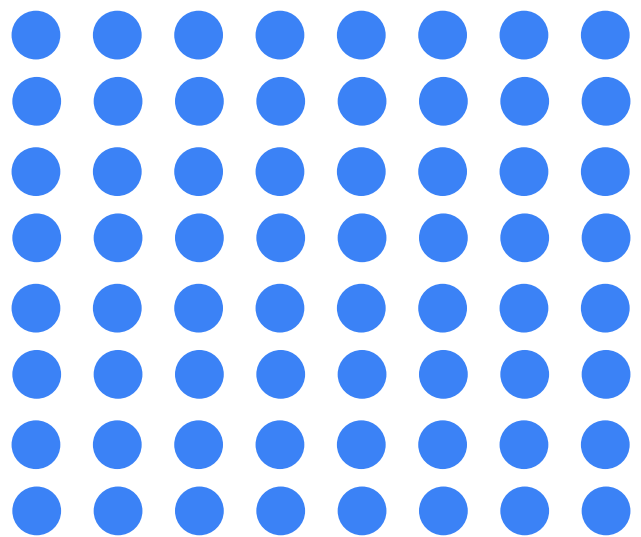


What is that?



A process for updating the model
and tracking your versions

Oh yeah



Formalizing Non-Written Protocols

Qualitative Subcomponents

Policies and Procedures

Assess the comprehensiveness of model-related policies and their implementation

Model Documentation

Evaluate the quality, completeness, and accessibility of model documentation

Change Management

Review processes for updating models and tracking model versions

Model Administration

Assess controls on access, usage, and roles/responsibilities

Next Steps

- 1 Workshop Session 2**
Quantitative aspects of model validation
- 2 Implementation Planning**
Validation schedules for Q1-Q4 2025
- 3 Resource Hub**
Access the Independent Validation Guidelines

Thank You