

Salesforce CRM Analytics and Einstein Discovery Consultant Certification Study Guide

Topic 6: Einstein Discovery Story Design (19% Exam Weight)

Introduction to Einstein Discovery Story Design

The **Einstein Discovery Story Design** topic, weighted at 19%, is a vital component of the Salesforce CRM Analytics and Einstein Discovery Consultant certification. This section focuses on leveraging Einstein Discovery—an AI-driven analytics tool within CRM Analytics—to uncover predictive insights, identify key drivers of outcomes, and provide actionable recommendations. It tests the ability to prepare data for analysis, create and interpret Einstein Discovery stories, and refine models to ensure accuracy and business relevance. Mastery of story design is essential because it transforms raw data into strategic insights, enabling organizations to predict trends, optimize processes, and make data-driven decisions.

Importance of Einstein Discovery Story Design

- **Predictive Power:** Identifies patterns and forecasts outcomes (e.g., “What drives customer churn?”), offering a competitive edge.
- **Actionable Insights:** Translates complex data into clear recommendations (e.g., “Increase support staffing to reduce resolution time”).
- **Business Impact:** Directly ties analytics to measurable goals like revenue growth, customer retention, or operational efficiency.
- **Automation:** Reduces manual analysis time, allowing rapid insight generation with AI. A poorly designed story—using flawed data or misinterpreting results—leads to inaccurate predictions, wasted resources, and missed opportunities, undermining trust in AI-driven analytics.

Exam Objectives for Story Design

While the Salesforce Exam Guide does not explicitly list sub-objectives, the scope of the Einstein Discovery Story Design topic implies the following key focus areas:

1. Prepare and structure data appropriately for Einstein Discovery analysis.
2. Create and analyze Einstein Discovery stories to extract meaningful insights and predictions.
3. Refine models and integrate results (e.g., writebacks to Salesforce) for practical application. This guide will explore these areas with exhaustive depth, providing a

comprehensive resource to excel in the 19% of the exam dedicated to Einstein Discovery Story Design.

Key Concepts and Subtopics: A Comprehensive Breakdown

The Story Design topic is divided into three critical areas: **Data Preparation for Stories**, **Story Creation and Analysis**, and **Model Refinement and Writebacks**. Each is dissected below with granular explanations, step-by-step configurations, extensive examples, practical scenarios, troubleshooting insights, and best practices to ensure a thorough understanding.

1. Data Preparation for Stories

- **Definition:** Data Preparation for Stories involves cleaning, structuring, and enriching datasets to meet Einstein Discovery’s requirements for AI analysis, ensuring the data is suitable for generating reliable insights.
- **Significance:** High-quality data is the foundation of accurate predictions—garbage in, garbage out applies here; incomplete or messy data skews results.
- **Mechanics:**
 - **Requirements:** Numeric outcome variable (e.g., “Revenue”), predictors (e.g., “Region,” “Product”), minimum 400 rows.
 - **Tools:** Recipes or Dataflows in CRM Analytics to transform data (e.g., filter, join, calculate fields).
- **Detailed Configuration Steps:**

1. Select Dataset:

- In Analytics Studio > Datasets, choose or create a dataset (e.g., “Sales Data” from Salesforce Opportunities).
- Verify: Contains outcome (e.g., “Amount”) and predictors (e.g., “Region,” “Close Date”).

2. Clean Data:

- Open Recipe in Data Manager > Recipes > New Recipe.
- Filter: Remove nulls (e.g., “Amount is not null”).

- Handle Outliers: Cap extreme values (e.g., “Amount > 1M → 1M” via formula).
- Standardize: Convert text to consistent formats (e.g., “USA” to “United States”).

3. **Enrich Data:**

- Add Calculated Fields: “Days to Close” (CloseDate - CreatedDate).
- Join: Merge with “Accounts” on “AccountId” for additional predictors (e.g., “Industry”).
- Aggregate: Ensure one row per analysis unit (e.g., one row per Opportunity).

4. **Validate:**

- Check row count (e.g., 5,000 rows > 400 minimum).
- Preview: Ensure numeric outcome (e.g., “Amount” as Number), predictors populated (e.g., “Region” has 5 values).
- Save as “Sales Analysis” dataset.

- **Customization Options:**

- **Buckets:** Group predictors (e.g., “Amount”: Low < \$10K, High > \$100K).
- **Normalization:** Scale numeric fields (e.g., “Amount” to 0-1 range via formula).

- **Troubleshooting Tips:**

- **Not Enough Rows:** Expand dataset (e.g., add historical data to reach 400+).
- **Missing Values:** Fill nulls (e.g., “Unknown” for blank “Region”) or exclude rows.

- **Best Practices:**

- Aim for 1,000+ rows for robust insights (400 is minimum, not optimal).
- Use meaningful predictors (e.g., “Industry” over “Created By”).
- Document prep steps (e.g., “Filtered null Amounts, joined Accounts”).

- **Practical Example:** A retailer preps “Sales Data”:

- **Scenario:** Analyze sales amount drivers.
- **Action:**
 - Dataset: “Opportunities” (Amount, Region, Product, Close Date).
 - Recipe: Filter “Stage = Closed Won,” add “Days to Close,” join “Product Catalog” for “Category.”
 - Output: “Sales Analysis” (10,000 rows).
- **Outcome:** Clean, enriched dataset ready for Einstein analysis.

2. Story Creation and Analysis

- **Definition:** Story Creation and Analysis involve configuring an Einstein Discovery story, running the AI analysis, and interpreting the resulting insights, predictions, and recommendations.
- **Significance:** This is where AI uncovers patterns (e.g., “Longer resolution times increase churn”) and provides actionable guidance, bridging data to decisions.
- **Mechanics:**
 - **Story Types:** Maximize (e.g., revenue), Minimize (e.g., churn), Predict (e.g., likelihood).
 - **Outputs:** Key drivers, correlations, predictions, “what-if” scenarios.
- **Detailed Configuration Steps:**
 1. **Create Story:**
 - In Analytics Studio > Create > Story.
 - Select dataset: “Sales Analysis.”
 - Define goal:
 - Outcome: “Amount” (Maximize).
 - Predictors: All fields (e.g., “Region,” “Days to Close,” “Category”) or exclude irrelevant (e.g., “Opportunity ID”).
 - Name: “Sales Maximization Story.”
 - Click “Create,” wait for analysis (e.g., 5-10 minutes for 10,000 rows).
 2. **Analyze Results:**

- **Overview:** View outcome stats (e.g., Avg Amount: \$50K, Range: \$1K-\$1M).
- **Key Drivers:** Top factors (e.g., “Days to Close” explains 35% of variance).
- **Correlations:** Visuals (e.g., bar chart: shorter “Days to Close” → higher Amount).
- **Predictions:** Model accuracy (e.g., $R^2 = 0.85$, 85% variance explained).
- **Recommendations:** Insights (e.g., “Reduce Days to Close by 5 days to increase Amount by 10%”).

3. Interpret Insights:

- Drill into drivers (e.g., “Category = Electronics” doubles Amount vs. “Clothing”).
- Test “what-if” (e.g., set “Days to Close” to 10, predict Amount jumps to \$60K).

4. Validate:

- Check model fit (e.g., $R^2 > 0.7$ for reliability).
- Cross-check with raw data (e.g., avg Amount aligns with dataset).

- **Customization Options:**

- **Exclude Predictors:** Remove noise (e.g., “Created By” if irrelevant).
- **Advanced Settings:** Adjust confidence intervals (e.g., 95% vs. 90%).

- **Troubleshooting Tips:**

- **Low Accuracy:** Add predictors (e.g., “Industry”), improve data quality.
- **No Insights:** Check outcome (e.g., “Amount” must be numeric, not text).

- **Best Practices:**

- Start with broad predictors, refine later (e.g., exclude weak drivers post-analysis).
- Focus on top 3-5 drivers for clarity (e.g., “Days to Close,” “Category”).

- Save story versions (e.g., “Sales_v1”) for iteration.
- **Practical Example:** A bank creates “Churn Reduction Story”:
 - **Scenario:** Minimize customer churn.
 - **Action:**
 - Dataset: “Customer Data” (Churn = Yes/No, predictors: Tenure, Support Calls).
 - Story: Minimize “Churn.”
 - Results: “Support Calls > 5” increases churn 40%, recommendation: “Reduce calls with proactive support.”
 - **Outcome:** Clear drivers and actions to retain customers.

3. Model Refinement and Writebacks

- **Definition:** Model Refinement and Writebacks involve improving the Einstein Discovery model’s accuracy and integrating predictions or recommendations back into Salesforce for operational use.
- **Significance:** Refinement ensures reliability, while writebacks make insights actionable (e.g., flag high-churn risks in Salesforce), closing the loop from analysis to execution.
- **Mechanics:**
 - **Refinement:** Adjust data, predictors, or settings to boost model performance.
 - **Writebacks:** Deploy predictions as fields (e.g., “Churn Probability”) or records in Salesforce.
- **Detailed Configuration Steps:**

1. Refine Model:

- Open story (e.g., “Sales Maximization Story”) > Edit.
- **Data:** Add predictors (e.g., “Lead Source”) via updated dataset.
- **Settings:** Exclude weak drivers (e.g., “Created Date” with <5% impact).
- Rerun analysis, compare R^2 (e.g., from 0.85 to 0.90).

2. **Deploy Model:**

- In Story > Deploy Model.
- Name: “Sales Prediction Model.”
- Options: Real-time (API) or batch (scheduled).
- Save and activate.

3. **Configure Writebacks:**

- In Deploy > Write Back to Salesforce.
- Map predictions:
 - Field: “Predicted Amount” (new custom field on Opportunity).
 - Object: Opportunity.
- Schedule: Daily batch update.
- Test: Run on 100 records, verify “Predicted Amount” populates (e.g., \$55K).

4. **Validate:**

- Check accuracy post-refinement (e.g., $R^2 > 0.85$).
- Confirm writebacks (e.g., Opportunity “Predicted Amount” matches story prediction).

• **Customization Options:**

- **Thresholds:** Write back only high-confidence predictions (e.g., >80% probability).
- **Flow Integration:** Use Flow to trigger actions (e.g., email on high churn risk).

• **Troubleshooting Tips:**

- **Model Weak:** Add rows (e.g., 500 → 2,000), refine predictors.
- **Writeback Fails:** Check field permissions (e.g., “Predicted Amount” not editable).

• **Best Practices:**

- Iterate models 2-3 times for optimal fit (e.g., test predictor impact).

- Use writebacks sparingly (e.g., 1-2 key fields) to avoid clutter.
 - Monitor deployed models monthly (e.g., re-run story if data shifts).
 - **Practical Example:** A tech firm refines “Support SLA Story”:
 - **Scenario:** Predict SLA breaches, write back risk scores.
 - **Action:**
 - Refine: Add “Agent Experience” predictor, improve R^2 from 0.80 to 0.87.
 - Deploy: “SLA Breach Model,” write “Breach Risk %” to Case object daily.
 - **Outcome:** Cases flagged (e.g., 75% risk), agents prioritize effectively.
-

Scenario Example: Comprehensive Story Design

Scenario: A retail chain needs an Einstein Discovery story to maximize sales revenue, identifying key drivers and writing predictions back to Opportunities.

- **Requirements Breakdown:**
 - **Goal:** Maximize “Amount” in “Opportunities” dataset.
 - **Data:** Salesforce Opportunities (Amount, Region, Product, Days to Close).
 - **Output:** Key drivers, predictions, “Predicted Amount” field.
- **Solution Design:**
 - **Data Preparation:**
 - **Dataset:** “Opportunities” (10,000 rows).
 - **Recipe:**
 - Filter: “Stage = Closed Won.”
 - Calculate: “Days to Close” (CloseDate - CreatedDate).
 - Join: “Products” on “ProductId” for “Category.”
 - Output: “Sales Analysis” (8,000 rows).

- **Validation:** “Amount” numeric, predictors populated (e.g., 5 Regions, 10 Categories).
- **Story Creation and Analysis:**
 - **Setup:**
 - Create Story: Analytics Studio > Story > “Sales Maximization Story.”
 - Goal: Maximize “Amount.”
 - Predictors: “Region,” “Days to Close,” “Category.”
 - **Results:**
 - Drivers: “Days to Close” (40%), “Category” (30%).
 - Insights: “Electronics” + “Days to Close < 15” → Amount +20%.
 - Prediction: $R^2 = 0.88$.
 - **Analysis:** “What-if” shows reducing “Days to Close” from 20 to 10 boosts Amount by \$10K avg.
- **Model Refinement and Writebacks:**
 - **Refine:**
 - Add “Lead Source,” rerun ($R^2 \rightarrow 0.90$).
 - **Deploy:**
 - Model: “Sales Prediction Model,” batch daily.
 - Writeback: “Predicted Amount” (Opportunity field), maps prediction (e.g., \$60K).
 - **Validation:** 100 Opportunities updated, “Predicted Amount” aligns with story.
- **Outcome:**
 - Key drivers: “Days to Close,” “Category.”
 - Recommendation: Shorten sales cycle for Electronics.
 - Sales team uses “Predicted Amount” to prioritize deals.

Exam-Focused Insights and Strategies

- **Common Questions:**
 - **Scenario-Based:** “Design a story to minimize churn with writebacks.” (Answer: Prep data with Churn outcome, analyze drivers, write risk % to Contacts.)
 - **Data Prep:** “What’s the minimum row count?” (Answer: 400, aim for 1,000+.)
 - **Troubleshooting:** “Model accuracy is low. Fix it.” (Answer: Add predictors, clean outliers.)
- **Key Memorization:**
 - Requirements: 400 rows, numeric outcome.
 - Story Outputs: Drivers, R^2 , recommendations.
 - Writeback Steps: Deploy, map fields, schedule.
- **Practical Tips:**
 - Build in Sandbox: Prep a 1,000-row dataset, run a story, deploy a prediction.
 - Know Limits: 100 predictors max, 10M rows per dataset.

Study Tips for Story Design

1. **Hands-On Practice:**
 - Prep a dataset (e.g., Cases), run a story (e.g., minimize resolution time).
 - Refine with new predictors, write back a prediction field.
2. **Memorize Concepts:**
 - Data Prep: Outcome vs. predictors.
 - Story Analysis: R^2 , key drivers.
3. **Scenario Mastery:**
 - Solve: “Maximize revenue, predict Opportunity values.”
4. **Trailhead Modules:**

- “Einstein Discovery Basics”
- “Build Stories with Einstein Discovery”

5. Test Edge Cases:

- Use a 300-row dataset (fails), fix with more data.
 - Deploy a weak model, refine for accuracy.
-

Summary of Story Design

This massive guide has delivered a comprehensive mastery of Einstein Discovery Story Design. It has covered:

- Preparing data with cleaning, enrichment, and validation for AI analysis.
- Creating and analyzing stories to uncover drivers, predictions, and recommendations.
- Refining models and implementing writebacks for actionable integration.

This exhaustive resource ensures readiness for the 19% of the exam focused on Story Design. This completes the six-topic series for the CRM Analytics and Einstein Discovery Consultant certification—proceed to review or additional practice as needed!