**From Deployment to Feedback**

When you complete this module, you’ll be able to describe the deployment and feedback stages of the data science methodology. You’ll learn how to assess a data model’s performance, impact, and readiness. You’ll be able to identify the stakeholders who usually contribute to model refinement. You’ll also be able to explain why deployment and feedback should be an iterative process. To complete your hands-on lab experience, you’ll devise a business problem to solve using data related to email, hospitals, or credit cards. You’ll demonstrate your understanding of data science methodology by applying it to a given problem. You’ll construct responses that address each phase of the CRISP-DM based on a chosen business problem. After submitting your work, you’ll evaluate your peers’ final projects and provide constructive ideas and suggestions that fellow learners can apply right away.

**Learning Objectives**

* Describe the deployment and feedback phases of the data science methodology.
* Determine when a model is ready to deploy.
* Devise a plan to elicit feedback from stakeholders involved in the data analysis process.
* Examine how feedback helps to refine a model.
* Assess the performance and impact of a data model.
* Explain why deployment and feedback should be an iterative process.
* Devise a business problem to be solved with data related to either email, hospitals, or credit cards.
* Demonstrate your understanding of data science methodology by applying it to a given problem. Construct responses that address each phase of the CRISP-DM based on a chosen business problem.
* Evaluate your peers’ final projects using the given rubric. Provide constructive feedback and offer ideas and suggestions that fellow learners can apply right away.

# **Deployment**

Simplified Explanation:

* **Deployment** is when a data science model is put into action to solve a real-world problem.
* It's important to help the people who will use the model understand how it works. This includes different roles like marketing, IT, and application developers.
* After the model is tested and confirmed to work, it can be released to a small group of users first to build confidence before a wider rollout.
* In a case study, a program was created to help doctors identify patients at risk of being readmitted to the hospital after discharge. The program needed to be easy to use and provide real-time risk assessments.
* Training for the clinical staff is essential, and collaboration with IT is needed to track patient outcomes and improve the model over time.

Summary:

* The video covers the **Deployment** stage of data science, emphasizing the importance of stakeholder engagement and training.
* It illustrates a case study where a model is used to assess patient readmission risks, highlighting the need for user-friendly applications and ongoing collaboration for effective implementation and feedback.

# **Feedback**

Simple Explanation:

* **Feedback is important**: After deploying a data model, it's crucial to gather feedback from users. This helps improve the model and assess how well it performs.
* **Cyclical process**: The methodology is cyclical, meaning each step leads to the next, allowing for continuous improvement.
* **Case study example**: In a case study about congestive heart failure, the feedback process includes:
  + Defining how to review results.
  + Tracking patients and their re-admission rates.
  + Measuring the effectiveness of the intervention.
* **Refinement**: Based on feedback, the model and intervention can be adjusted to improve outcomes. This may include adding new data or refining processes.

Summary:

The video emphasizes the significance of the feedback stage in the data science methodology. It outlines how feedback helps refine models and improve their effectiveness over time, using a case study on congestive heart failure as an example. The process involves tracking patient outcomes, measuring intervention success, and making necessary adjustments based on the collected data.

# **Storrytelling**

Simple Explanation:

* **Storytelling with Data**: Data analysts need to tell clear and engaging stories using data. This helps others understand the information better.
* **Importance**: If you can't explain your data well, it won't matter how good your findings are. You need to connect with your audience emotionally.
* **Balance**: It's important to keep the story simple while also including necessary details from the data.
* **Research Findings**: A study showed that people remember stories better than just numbers. So, telling a story helps make your message stick.

Summary:

The video emphasizes that storytelling is a crucial skill for data analysts. It helps them communicate their findings effectively, ensuring that the audience understands and remembers the information. The ability to create a compelling narrative around data can drive action and engagement, making it essential for a successful career in data analysis.

# **Course Summary**

Key Points:

* **Data Science Thinking**: You learned how to approach data science problems methodically.
* **Steps Involved**:
  + **Identify the Problem**: Formulate a clear business or research question.
  + **Data Collection**: Gather relevant data for analysis.
  + **Data Analysis**: Analyze the data to find insights.
  + **Model Building**: Create a model based on the data.
  + **Feedback and Improvement**: After deploying the model, gather feedback to enhance it.

Summary:

The course taught you the steps to think like a data scientist, including how to define problems, collect and analyze data, build models, and use feedback for improvement. You also learned that these steps are iterative, meaning you can always refine your model as new data becomes available. A real case study illustrated how these methodologies can improve business outcomes, such as enhancing patient care in healthcare settings.

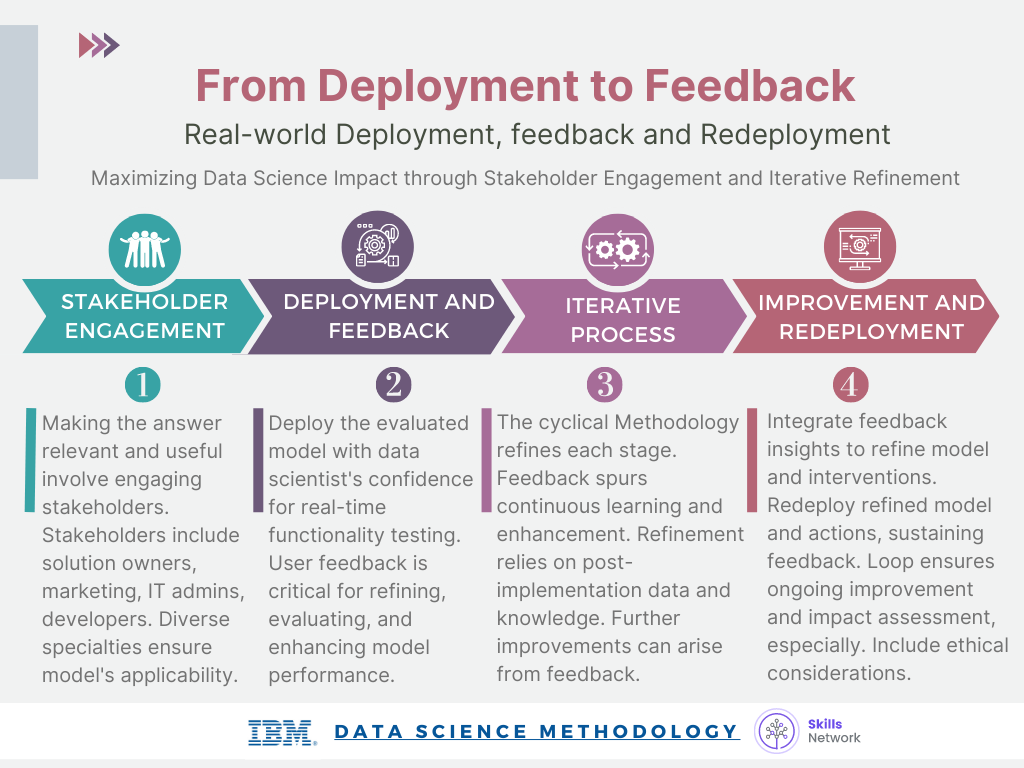
# **Lesson Summary**

**Lesson summary**

**Module 3 Lesson 1: Deployment to Feedback**

Congratulations! You have completed this lesson. At this point in the course, you know:

* Stakeholders, including the solution owner, marketing staff, application developers, and IT administration evaluate the model and contribute feedback.
* During the Deployment stage, data scientists release the data model to a targeted group of stakeholders.
* Stakeholder and user feedback help assess the model's performance and impact during the Feedback stage.
* The model's value depends on iteration; that is, how successfully the data model incorporates user feedback.



# **Glossary**

**Glossary: From Deployment to Feedback**

Welcome! This alphabetized glossary contains many of the terms you'll find within this lesson. These terms are important for you to recognize when working in the industry, when participating in user groups, and when participating in other certificate programs.

|  |  |
| --- | --- |
| **Term** | **Definition** |
| **Browser-based application** | An application that users access through a web browser, typically on a tablet or other mobile device, to provide easy access to the model's insights. |
| **Cyclical methodology** | An iterative approach to the data science process, where each stage informs and refines the subsequent stages. |
| **Data collection refinement** | The process of obtaining additional data elements or information to improve the model's performance. |
| **Data science model** | The result of data analysis and modeling that provides answers to specific questions or problems. |
| **Feedback** | The process of obtaining input and comments from users and stakeholders to refine and improve the data science model. |
| **Model refinement** | The process of adjusting and improving the data science model based on user feedback and real-world performance. |
| **Redeployment** | The process of implementing a refined model and intervention actions after incorporating feedback and improvements. |
| **Review process** | The systematic assessment and evaluation of the data science model's performance and impact. |
| **Solution deployment** | The process of implementing and integrating the data science model into the business or organizational workflow. |
| **Solution owner** | The individual or team responsible for overseeing the deployment and management of the data science solution. |
| **Stakeholders** | Individuals or groups with a vested interest in the data science model's outcome and its practical application, such as solution owners, marketing, application developers, and IT administration. |
| **Storytelling** | Storytelling is the art of conveying your message, or ideas through a narrative structure that engages, entertains, and resonates with the audience. |
| **Test environment** | A controlled setting where the data science model is evaluated and refined before full-scale implementation. |

**Final Poject**

# **Introduction to CRISP-DM**

**CRISP-DM**, which stands for **Cross-Industry Standard Process for Data Mining**. Here’s a simple explanation and summary:

Explanation:

* **CRISP-DM** is a method used by data scientists to guide their work in data mining projects.
* It has **six main steps** that help in solving data problems:
  1. **Business Understanding**: Understand what the business needs and what problem you are trying to solve.
  2. **Data Understanding**: Collect and explore the data you will use.
  3. **Data Preparation**: Clean and organize the data so it can be used effectively.
  4. **Modeling**: Create models to analyze the data and find patterns.
  5. **Evaluation**: Test the models to see if they work well and meet the business goals.
  6. **Deployment**: Use the model on new data and make any necessary changes.

Summary:

* **CRISP-DM** is an iterative process, meaning you might need to revisit earlier steps based on what you learn.
* Communication with stakeholders is crucial throughout the process to ensure everyone is on the same page.
* After completing all six steps, a final meeting is held to review the results and ensure they address the business problem.

# **Review What you learned?**

**Review what you learned**

After completing this course, you learned many facts about data science methodology. Here are 14 key, high-level takeaway facts you’ll want to remember from this course.

* Foundational methodology, a cyclical, iterative data science methodology developed by John Rollins, consists of 10 stages, starting with Business Understanding and ending with Feedback.
* CRISP-DM, an open source data methodology, combines several data-related methodology stages into one stage and omits the Feedback stage resulting in a six-stage data methodology.
* The primary goal of the Business Understanding stage is to understand the business problem and determine the data needed to answer the core business question.
* During the Analytic Approach stage, you can choose from descriptive diagnostic, predictive, and prescriptive analytic approaches and whether to use machine learning techniques.
* During the Data Requirements stage, scientists identify the correct and necessary data content, formats, and sources needed for the specific analytical approach.
* During the Data Collection stage, expert data scientists revise data requirements and make critical decisions regarding the quantity and quality of data. Data scientists apply descriptive statistics and visualization techniques to thoroughly assess the content, quality, and initial insights gained from the collected data, identify gaps, and determine if new data is needed, or if they should substitute existing data.
* The Data Understanding stage encompasses all activities related to constructing the data set. This stage answers the question of whether the collected data represents the data needed to solve the business problem. Data scientists might use descriptive statistics, predictive statistics, or both.
* Data scientists commonly apply Hurst, univariates, and statistics such as mean, median, minimum, maximum, standard deviation, pairwise correlation, and histograms.
* During the Data Preparation stage, data scientists must address missing or invalid values, remove duplicates, and validate that the data is properly formatted. Feature engineering and text analysis are key techniques data scientists apply to validate and analyze data during the Data Preparation stage.
* The end goal of the Modeling stage is that the data model answers the business question. During the Modeling stage, data scientists use a training data set. Data scientists test multiple algorithms on the training set data to determine whether the variables are required and whether the data supports answering the business question. The outcome of those models is either descriptive or predictive.
* The Evaluation stage consists of two phases, the diagnostic measures phase, and the statistical significance phase. Data scientists and others assess the quality of the model and determine if the model answers the initial Business Understanding question or if the data model needs adjustment.
* During the Deployment stage, data scientists release the data model to a targeted group of stakeholders, including solution owners, marketing staff, application developers, and IT administration.,
* During the Feedback stage, stakeholders and users evaluate the model and contribute feedback to assess the model’s performance.
* The data model’s value depends on its ability to iterate; that is, how successfully the data model incorporates user feedback.