



Welcome to Vectice

Tutorial Instructions

Our tutorial project simulates a real retailer data science initiative to forecast in-store unit sales.

Before you start, we recommend you watch our 6-minute video introduction to get familiar with our web app. You can access it from <https://docs.vectice.com/introduction>.

In this tutorial, you will learn how to:

- Guide your team using phases, steps, and iterations.
- Use just a few lines of code to leverage Vectice Python APIs to log your key models, datasets, and code.
- Auto-document your work in Vectice.

Use case:

You are a data scientist working on a Linear Regression model alongside your retail team. The project leader added you to a project and provided you with a set of phases and underlying steps to complete.

As part of the work has been completed already, you are only expected to log your model to Vectice and document the related evaluation.

Required Resources:

To complete this tutorial, you will need the following resources:

- 1) Retail Project Example
 - This project is available from your personal workspace in your Vectice instance.
- 2) Modeling notebook
 - [The notebook is available on our GitHub repo](#)

Get started:

This tutorial provides two paths to completion:

- **Getting familiar with the Vectice web app:** In this scenario, we recommend you follow this Tutorial phase by phase. [Go to page 2](#)
- **Getting familiar with the Vectice API:** In this scenario, we recommend you jump right away into the notebook. [Go to page 3](#)

Contact us to learn more

Do not hesitate to contact support@vectice.com for any questions. Our API documentation and general Vectice documentation are also available.



Getting Familiar with Vectice

Once you have accessed your personal workspace and the project called "Tutorial: Tutorial Project: Forecast in-store unit sales (23.1)", on the project overview, you will notice phases that have been predefined by your managers. These phases follow the CRISP-DM framework, but your manager could have used any other framework!

Certain phases are already marked as "Completed". You do not have anything to do with these.

However, the Modeling phase is "In Progress." You have some work to complete before moving to the grey phases you have "Not Started."

The purpose of this path is to get you familiarized with our UI, as such, the content has already been provided as part of this tutorial, and no activity is required. This content is available as a reference only.

Business understanding phase

This phase describes the business problem, the goals, and the success criteria of the project. Optionally you can review this phase to better understand the use case; however, please feel free to proceed to the next phase, "Data Understanding".

Data understanding phase

In this phase, data scientists present the data and leverage graphical representations to describe it. Multiple iterations were completed for that phase, and iteration two was selected before moving on to the next phase. If you would like to repeat this section on your own, the supporting notebooks and datasets are available below:

- [Link to Data Understanding notebook in Colab](#)
- [Link to datasets in GitHub](#)

Data preparation phase

In this phase, data scientists create the cleaned dataset used in the modeling phase. One iteration was selected before the team moved on to the next phase. If you would like to repeat this section on your own, the supporting notebooks and datasets are available below:

- [Link to Data Preparation notebook in Colab](#)
- [Link to Datasets in GitHub](#)

Modeling phases

In this phase, the data scientist will explore model technics, splitting strategies, and other model architectures to generate a model that will satisfy the metrics and the goals defined by the business requirements.

At this point, you may also want to learn how to create a new iteration, follow the steps specified for this phase, log your first model version using the Vectice API, and easily document it. To do so, refer to page X.

You may also want to just look at the content and move on to the last phase of this approach.



Evaluation phase

In this phase, you will capture the results of your model and compare it to the business problem, the goals, and the success criteria of the project you defined during the business understanding phase. You would also evaluate the accuracy and performance of your model.

This phase of the project has been prepopulated for you. Feel free to assign ownership to yourself and add content for the model based on what you learned previously in the tutorial.

Deployment phase

In the sections, you would describe your deployment strategy. Deployment options depend on your infrastructure, model risk management, and MLOps processes.

This phase of the project has not been started but feel free to assign ownership to yourself and add content based on what you learned previously in the tutorial.

Getting familiar with the Vectice API

In this approach, you will learn how to create a new iteration, follow the steps specified for this phase and log your first model version using the Vectice API. You will also learn how to leverage the documentation features of Vectice to collaborate with your colleagues and quickly report on your findings. The phase content has already been provided.

Modeling

This section represents a typical data modeling phase. You are expected to mostly work from the notebook provided below and then look at the result in Vectice. Here are the steps you need to complete :

1. Click: [Link to Modeling Notebook](#) to open the notebook in **Google Colab**
 - or download the notebook from this [Link to Github folder](#) and run it in other notebook providers such as JupyterLab.
2. **Run the notebook and see the new iteration magic in your "Modeling" phase**
 - You can run the notebook cell by cell and see the magic along your changes, or you can run the entire notebook and see the final result on the iteration page.
3. **Insert an Iteration Widget:** In the evaluation phase, under the documentation section, at the top right corner, select Insert, then click on Iterations. The widget has been added at the bottom of the page.
4. **Display the latest Iteration:** Click on the widget, then click on Iteration 2, and then click on Insert. You should now see the details related to the latest iteration.
5. **Mark this phase status as "Completed"** to indicate you are done with this phase.
6. **Move to the "Evaluation" Phase**

What have you learned?

In this brief tutorial, as a Data Scientist, you have registered datasets and models and commented on the project directly within your code, providing context and insights into the project in real-time.

As a project lead, you have learned how to create a project based on your internal standard and processes and provide the Data Scientist with a blueprint of tasks needed.

This is the end of this tutorial. Feel free to explore the rest of the documentation capabilities on your own. For example, the history icon on the top right allows you to go back to previous versions and restore them. You can also @mention your colleagues to include them in conversations.