Organon Analytics Machine Learning Course Coding Preparation Document

Table of Contents

[About 2](#_Toc27472563)

[Visual Studio 2019 Installation 2](#_Toc27472564)

[GitHub Tutorial 5](#_Toc27472565)

[PostgreSQL Installation 5](#_Toc27472566)

[DBeaver Installation 10](#_Toc27472567)

[Connect to the Database in DBeaver 10](#_Toc27472568)

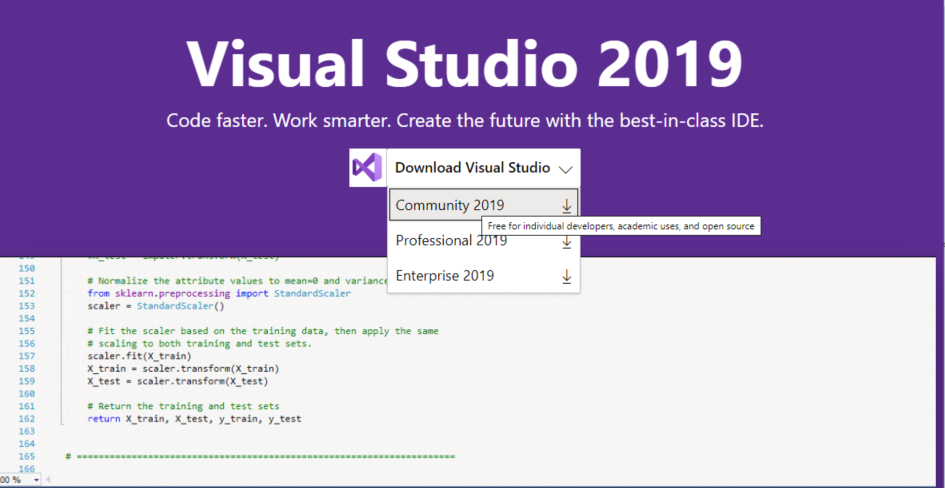
[Importing Sample Data 14](#_Toc27472569)

# About

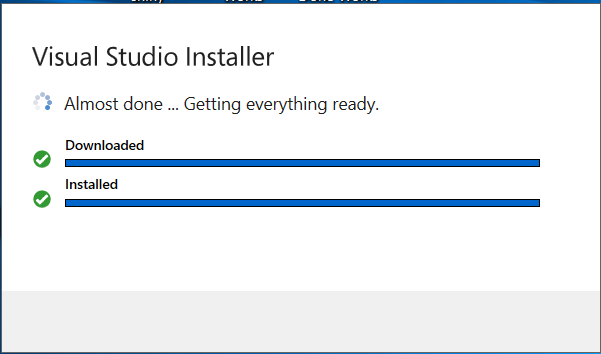
In this document we will introduce some tools you need to be familiar for developing GBM algorithm. As a development environment we will use Visual Studio 2019 and store out code on GitHub. The language for coding is C#. When developing GBM algorithm we will use PostgreSQL database as data source.

# Visual Studio 2019 Installation

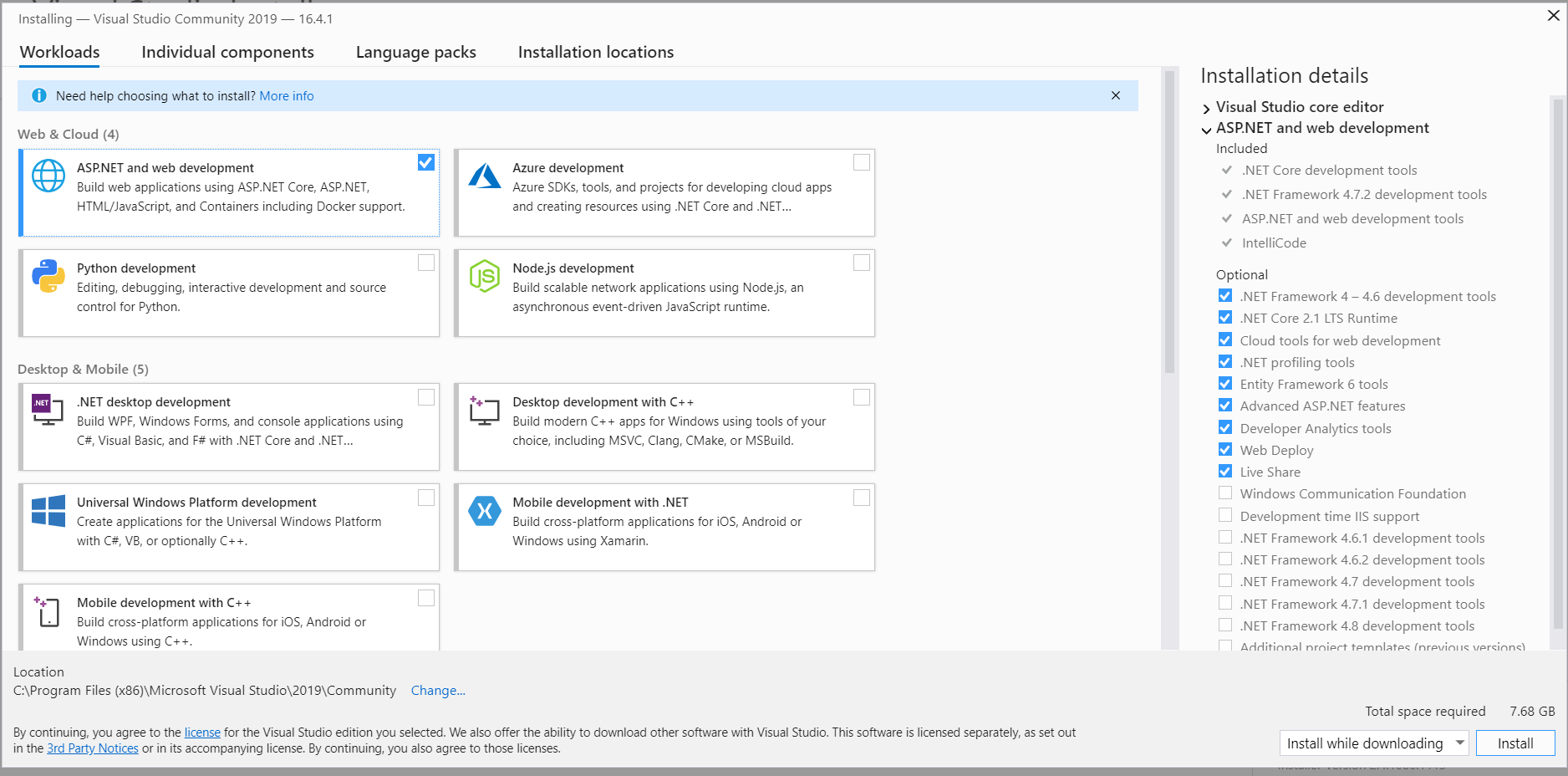
1. Please download Visual Studio 2019 Community Edition: <https://visualstudio.microsoft.com/vs/>



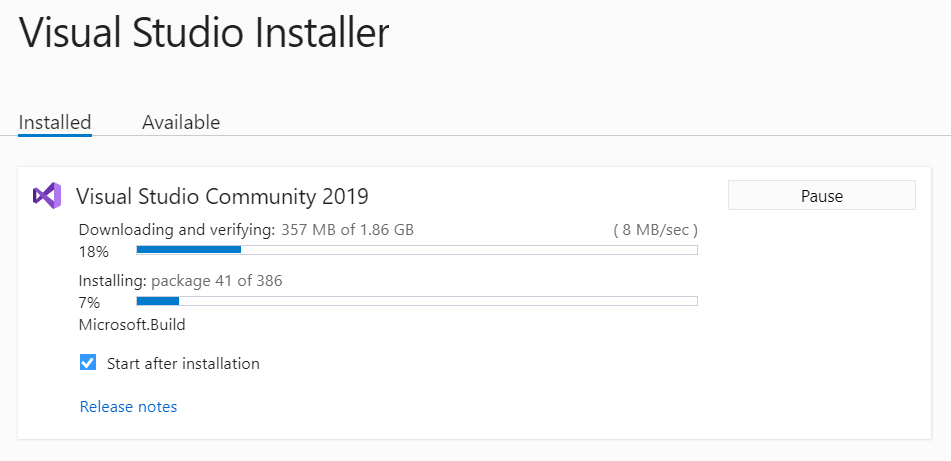
1. Start the Installer and click continue:



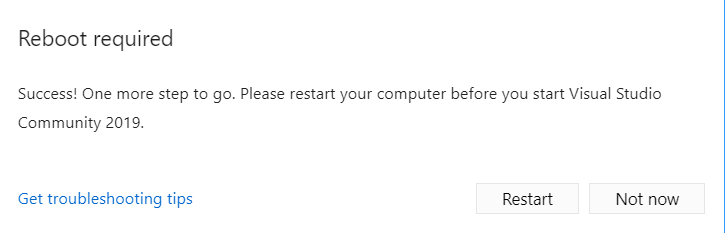
1. After installer is installed, you will see Workloads to select. Select “ASP.NET and web development” and click on Install.



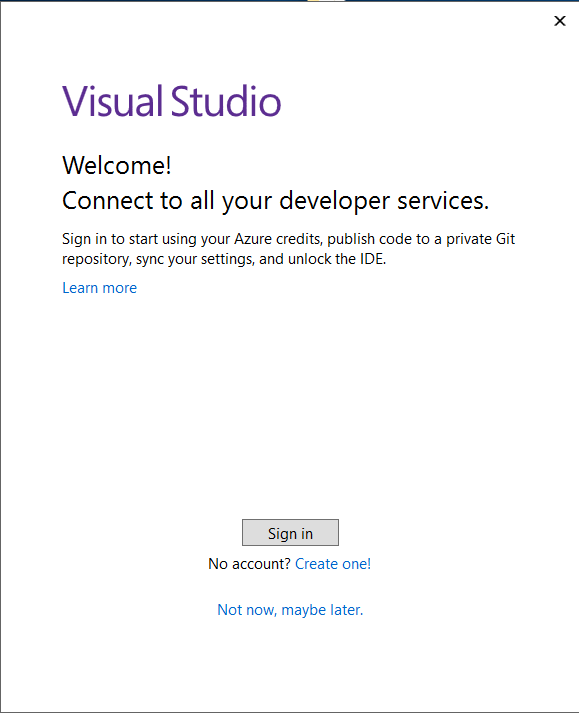
1. It will download and install components. Make sure you have a good internet connection. It will download nearly 2GB.



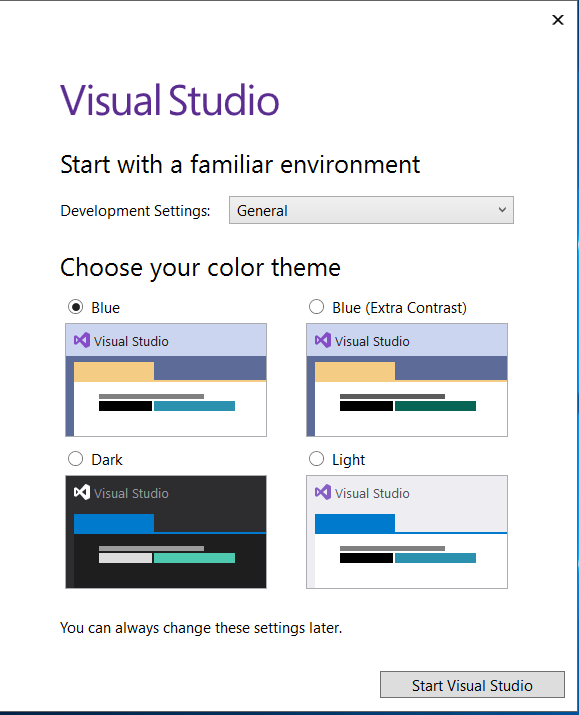
1. When installation ends restart your computer:



1. After restart start Visual Studio 2019 Application. You may create a developer account and Sign in. If you do not prefer, click on “Not now, maybe later”.



1. Choose your favourite theme and click “Start Visual Studio”



1. After some time it will start. You may create a new project or you can open GBM algorithm.

# GitHub Tutorial

We will develop and push GBM algorithm on our GitHub repository <https://github.com/Organon-Analytics/GBM>. Keep in mind that we cannot implement all of the algorithm in the class so there will be some coding assignments. You can fork our repository and make changes to existing code.

In the link you can find a tutorial for introduction to GitHub. <https://product.hubspot.com/blog/git-and-github-tutorial-for-beginners>

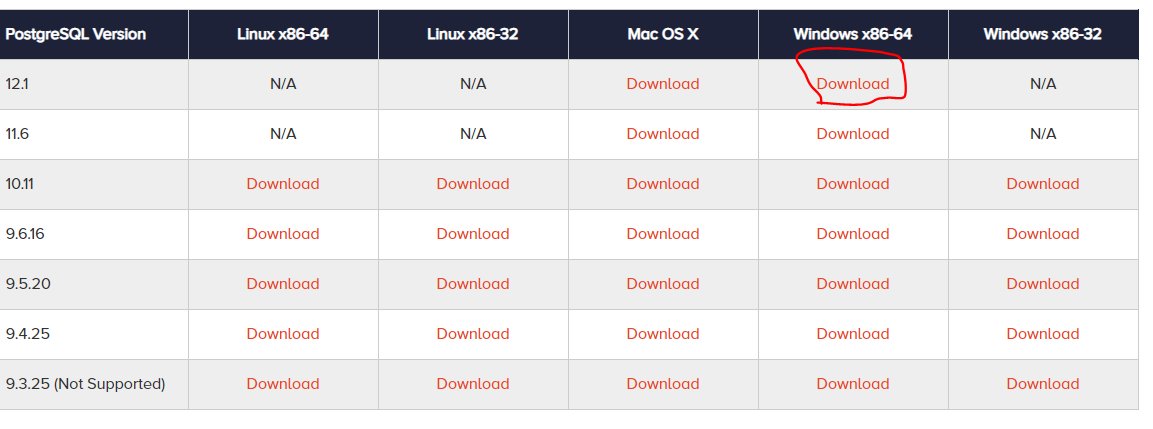
Most common git commands you will use are pull, push and commit. Be familiar with them. For the sake of simplicity you can use GitHub Desktop application <https://desktop.github.com/>.

# PostgreSQL Installation

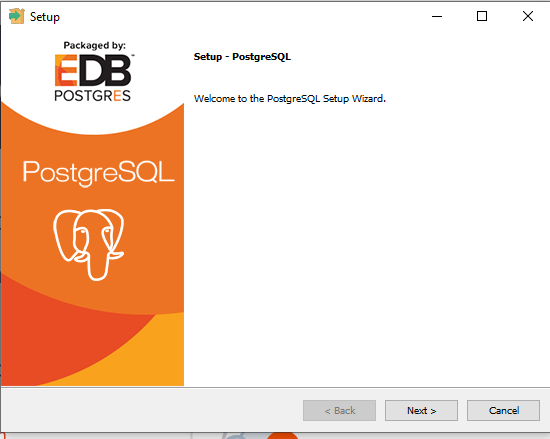
We will use PostgreSQL as the database. We will store our sample data and application outputs.

1. Go to <https://www.enterprisedb.com/downloads/postgres-postgresql-downloads>.

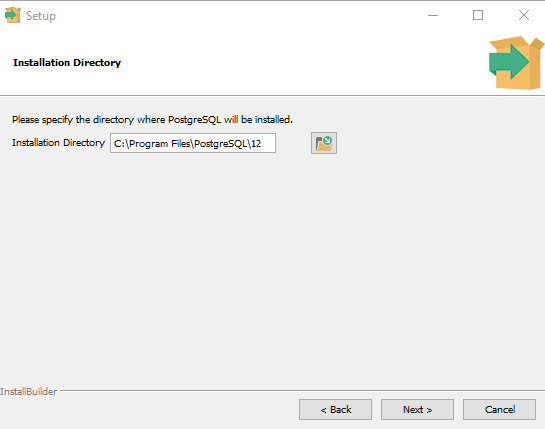
Download PostgreSql for you operating system. Next Steps wil be for Windows Operating system. We will install version 12.1 .



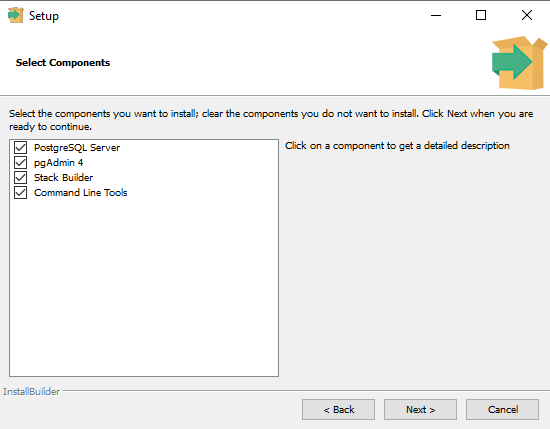
1. Start the installer and click on “Next”.



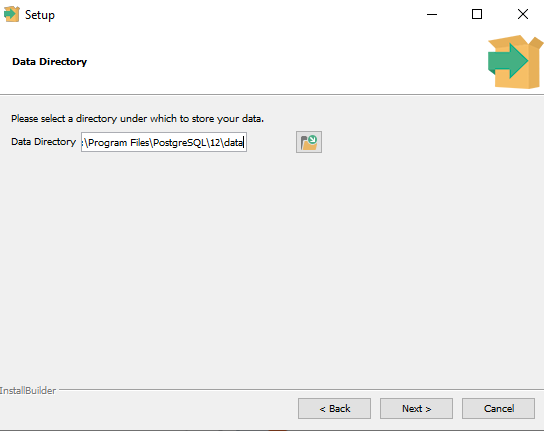
1. Choose Installation Directory and click on “Next”



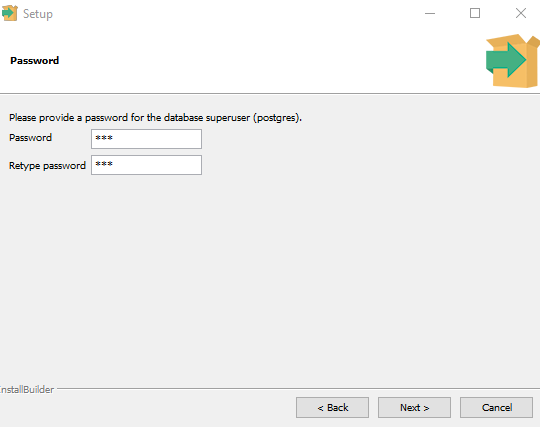
1. Select All components and click on “Next”



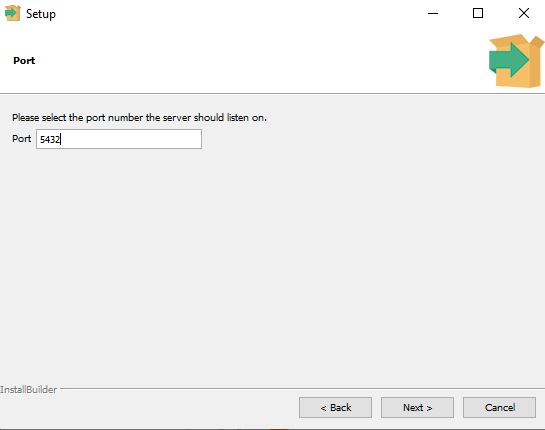
1. Choose Data Directory and click on “Next”



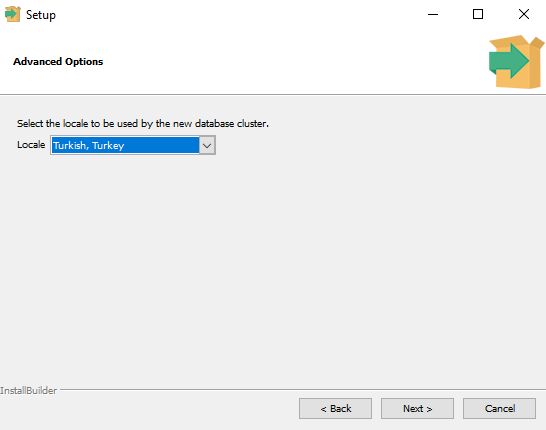
1. Set a password for “superuser” and click on “Next”



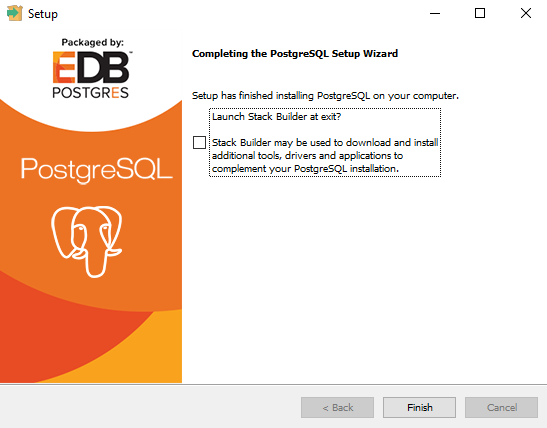
1. Choose a port. Default port is 5432. We will use 5432, if you are not sure about whether a port is in use or not, please leave it as 5432



1. Choose Turkey as locale



1. Click next and start installation.
2. Finish the installer



# DBeaver Installation

We will use DBevaer as a SQL client. We will run and get results of our SQL queries for development.

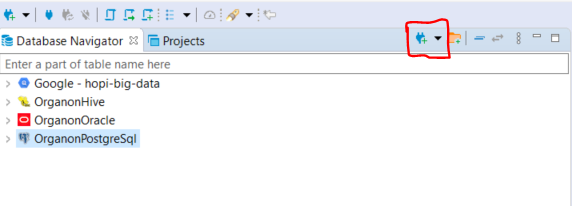
1. Download DBeaver <https://dbeaver.io/download/>. We will install on Windows



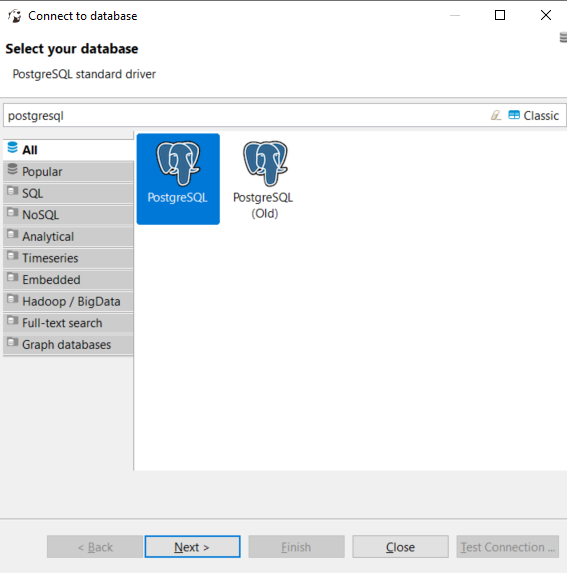
1. Start the installer
2. Do not change any configurations, just press on Next ☺

# Connect to the Database in DBeaver

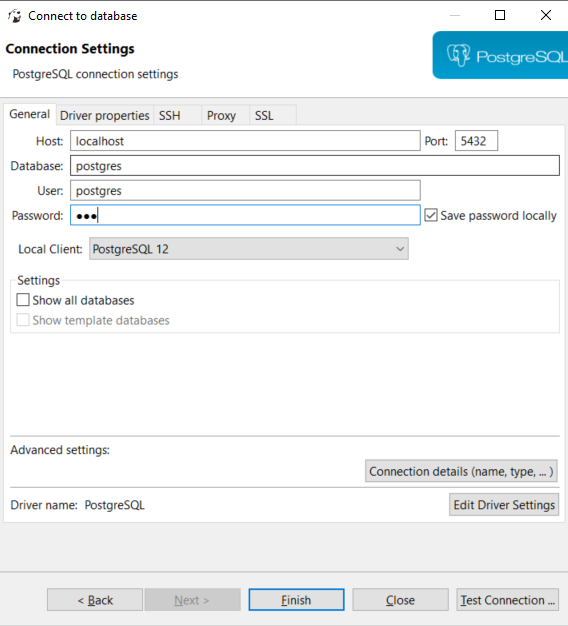
1. Open DBeaver after installation
2. Click on new Connection button



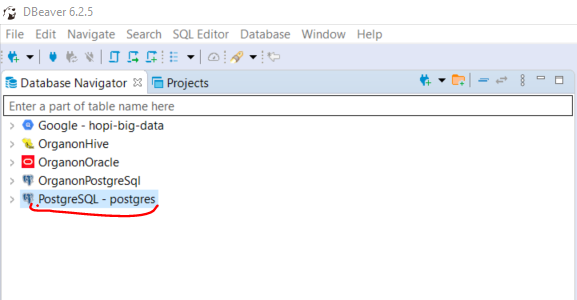
1. Search for “postgresql” and select “PostgreSQL” and click next



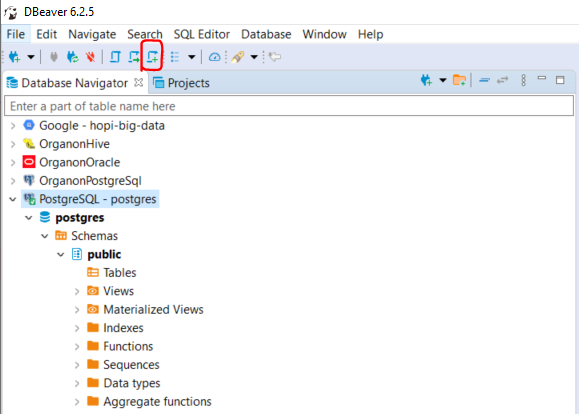
1. Enter the password you choose in PostgreSQL installation and click on “Test Connection”. It may want you to download some libraries. Download these libraries, they are for the connection. If connection is valid then click on “Finish”.



1. Righ Click on your Connection and choose “Connect”.



1. You can create new Sql Editor and work on it



# Importing Sample Data

1. Download “adult\_train\_combined\_postgresql.sql” file from <https://github.com/Organon-Analytics/ML-Course/tree/master/Data>
2. Open “adult\_train\_combined\_postgresql.sql” file with dbeaver. Select whole data (ctrl + A) and Execute script using “Execute SQL Script” button. It may take several minutes (nearly 30 minutes).

