# DAANY – .NET DAta ANalYtics library

A close up of a sign

Description automatically generated

## Introduction

Daany is .NET data analytics library written in C# language and it supposed to be a tool for data preparation, feature engineering and other kinds of data transformations prior to creating ml-ready data set. It is .NET based library and can be run on Windows Linux based distribution and Mac. It follows the .NET Standard 2.0.

Besides data analysis, the library implements a set of statistics or data science features e.g. time series decompositions, optimization performance parameters and similar.

The Daany project consists of four main components:

* Daany.DataFrame and
* Daany.Stats
* Daany.Math
* Daany.DataFrame.Ext

The main Daany component is Daany.DataFrame the DataFrame implementation for data analysis. It is much like Pandas but this component is not going to follow pandas implementation. Danny.DataFrame is suitable for doing data exploration and preparation with C# notebook. Daany.DataFrame doesn’t require any predefined class type in order to create or load data to DataFrame. All data are parsed internally in order to defined relevant value type of each column. The Daany.DataFrame implements set of powerful features for data manipulation, handling missing values, calculated columns, merging two or more data frame into one, and similar. It is handy for extracting its rows or columns as series of elements and put into the chart to visualizing the data.

Daany.Stat is a collection of statistics features e.g. time series decompositions, optimization, performance parameters and similar.

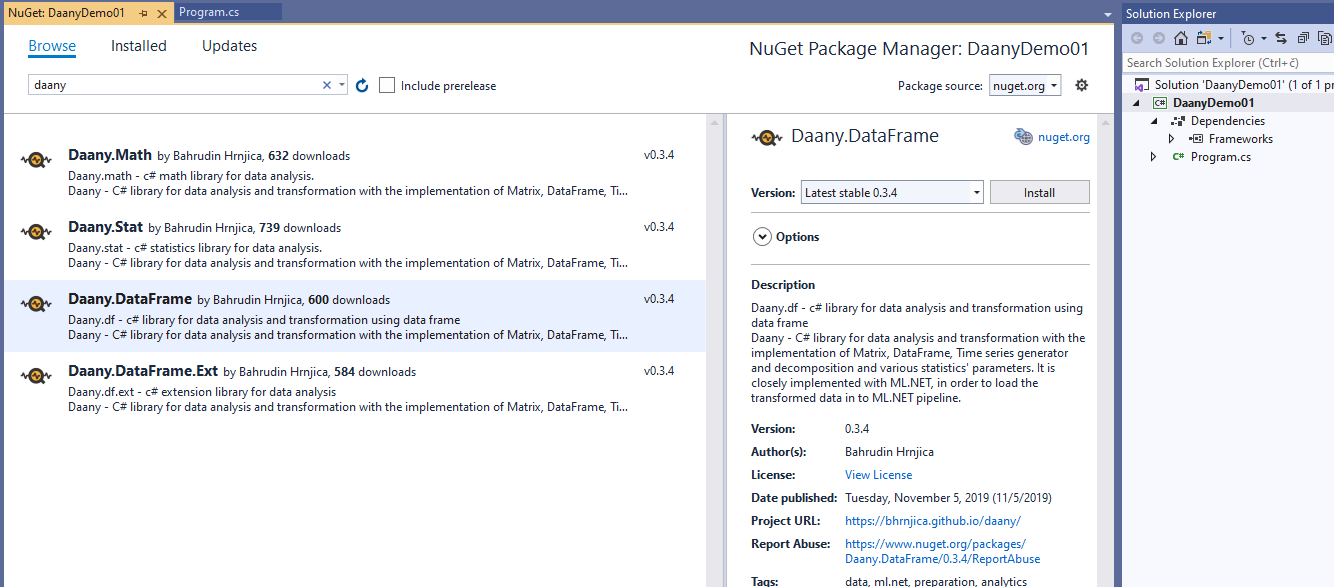
Daany.Math is a component with the implementation of od matrix and related linear algebra capabilities. It contains some implementation of other great open source projects. Daany.DataFrame.Ext contains extensions of DataFrame which are related to other projects mostly to ML.NET. The Daany.DataFrame should not be dependent on Ml.NET and other libraries. So any future DataFrame feature which depends on something other than Daany.Math, should be placed in Daany.Ext.

The project is developed as a need to have a set of data transformation features in one library while I work with machine learning. So I thought it might help others in the same situation. Currently, the library has pretty much data transformation features and might be your number one data transformation library on .NET platform.

# How to start with Daany

Daany is 100% .NET core and can be run on any platform .NET Core supports, from the Windows x86/x64 to Mac or Linux based OS. It can be used by Visual Studio or Visual Studio Code. Daany is a .NET component consisted of 4 NuGet packages, so the easiest way to start with it is to install the packages into your .NET application. Create or open your .NET application and open NuGet packages window within Visual Studio. Type Daany in browse edit box and hit enter. You can find four packages starting with Daany. You have few options to install the packages.

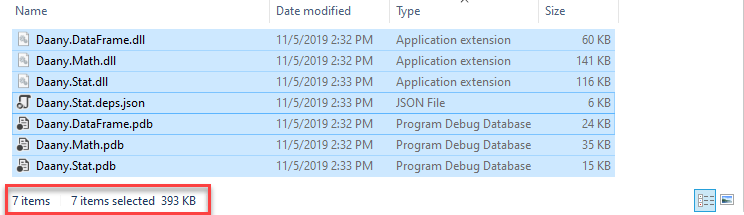
1. Install Daany.DataFrame – only. Use this option if you want only data analysis by using DataFrame. Once you click Install button, Daany.DataFrame and Daany.Math will be installed into your project app.
2. Install Daany.Stat package with time series decomposition and related statistics features.



Once you install the packages, you can start developing your app using Daany packages.

## Using Daany as assembly reference

Since Daany has no dependency on other library you can copy three dlls and add them as reference into your project.



In order to do so clone the project from <http://github.com/bhrnjica/daany> build it and copy Daany.DataFrame.dll, Daany.Math.dll and Daany.Stat.dll to your project as assembly references. Whole project is just 400 KB which is very small and handy. Currently Daany.DataFrame.Ext depends on ML.NET, and should be install as well.

## Namespaces in Daany

Daany project contains several namespaces for separating different implementation. The following list contains relevant namespaces:

* using Daany – data frame and related code implementation,
* using Daany.Ext – data frame extensions, used with dependency on third party library,
* using Daany.MathExt – math related stuff implemented in Daany,
* using Daany.Optimizers – set of optimizers like SGD,
* using Daany.Stat – set of statistics implementations in the project.