

# timeline Documentation

Manuel Haid

Thomas Mauerhofer

Matthias Wölbitsch

June 17, 2016

## Contents

### 1 Introduction

In this section we want to explain how to start our project.

- **go to the vagrant folder:** `cd vagrant/`
- **start vagrant:** `vagrant up` (the first time can take a long time, because of the new downloaded and installed dependencies)
- **start our website:** start a browser and go to `http://localhost:5000/`

Now our website should start and you can explore it.

### 2 Requirements

- `>= Python 3.5`

### 3 Dependencies

This section contains a description of all used external dependencies (e.g. libraries) and why they were used in this project. The python dependencies can be installed using `pip`, the Python package manager, using the `requirements.txt` file. It is recommended to use a virtual environment when developing for this project to avoid conflicts with other system-wide installed versions of the same dependencies.

#### 3.1 Back-end (Python)

**Flask** Flask is a microframework (simple but extendable) for web development. It allows RESTful request dispatching and is well documented. In this project Flask is used for the rendering of the front-end using its template engine and providing an API for the supported visualizations.

**Pandas** Pandas is a widely used data analysis library. It provides data structures to efficiently store and query in-memory data. All temporal data (i.e. all time series) in this project is stored in Pandas DataFrame or Series objects.

**NumPy** NumPy is a powerful package for scientific computing. It provides efficient data structures and algorithms to operate on them. However, in this project this library is only used to sample random data for the example data sets and is therefore not very important.

## 3.2 Front-end (HTML5)

**jQuery** jQuery is the feature rich JavaScript library to simplify client-side scripting. In this project it is mainly used for AJAX calls.

**MetricsGraphics.js** MetricsGraphics.js is used to render the time series visualizations. It is based on the well-known D3.js graphic library and provides all necessary primitives and features to plot the required figures.

**Bootstrap** This CSS framework is used to for its beautiful design templates to provide a familiar and responsive user interface. Furthermore, some bootstrap add-ons were used for additional widgets (e.g. the bootstrap datetime picker).

## 4 Modules

### 4.1 APP

This module simply creates the Flask object and registers the routes for the front-end and for the API. The Flask object must be used to start timeline using the `run` method of the app object. The source code is located in the `app.py` file.

### 4.2 API

The task of this module is to take requests from the front-end in form of AJAX calls and return the visualization primitives for the MetricsGraphics.js library as JSON objects. The source code is located in the `api.py` file.

**Time Series Plot:** `api/time_plot/{time_series_id}`

**Method:** GET

**Description:** This route create the visualization primitive for a simple time plot (i.e. a time plot of only one time series). It takes the identifier of the time series as part of the URL (e.g. `api/time_plot/42`) and can take additional parameter. It may return a HTTP status code 404 if no time series with the given identifier exists and a HTTP status code 400 for invalid values for the optional GET parameter.

**Additional Parameter:** All of the following parameter are optional. It is possible to only retrieve a slice of the time series using the `start_date` and/or `end_date` parameter.

Each of this parameter must be an integer, which represents to UNIX time stamp of the start time of the slice and the end time of the slice, respectively. Additional it is possible to add the graphs of the rolling mean and/or the rolling standard deviation to the plot. The `rolling_mean_window` and `rolling_std.window` parameter must be integer that represent the number of samples that is used to calculate the running mean and standard deviation, respectively.

#### **Time Series Plots:** `api/time_plots`

**Method:** GET

**Parameter:** list of time series ids

**Description:** This route create the visualization primitive for a simple time plot (i.e. a time plot of all time series). It may return a HTTP status code 404 if no time series with the given identifier exists.

#### **Live Time Series Plot:** `api/live_plot/{live_time_series_id}`

**Method:** GET

**Parameter:** last received

**Description:** This route create the visualization primitive for a simple live time plot. It may return a HTTP status code 404 if no time series with the given identifier exists. The `last_received` parameter must be integer that represent the last received timestamp, the start date and the new data.

#### **ACF Plot:** `api/acf_plot/{time_series_id}`

**Method:** GET

**Parameter:** max lag, scale

**Description:** This route create the visualization primitive for a simple acf time plot. It may return a HTTP status code 404 if no time series with the given identifier exists and a HTTP status code 400 for invalid values for the optional GET parameter.

**Additional Parameter:** All of the following parameter are optional. It is possible to only retrieve a slice of the time series using the `max_lag` parameter. The parameter must be an integer.

#### **Forecasting Plot:** `api/forecasting_plot/{forecast_id}`

**Method:** GET

**Parameter:** none

**Description:** This route create the visualization primitive for a simple forecasting time plot. It may return a HTTP status code 404 if no time series with the given identifier exists.

## 4.3 DATA SETS

The task of this module is to get the values for the time series, live data and forecast data and store it in their classes. The source code is located in the `data_sets.py` file.

### Register Data Sets:

**Method:** ADD

**Description:** In this section, we have three functions to add the data set to the `time_series`, `live_time_series` or `forecast` dictionary.

### Get Data Sets:

**Method:** GET

**Description:** In this section, we have different functions to get the data sets (`time_series`, `live_time_series` or `forecast`). There are functions to get one data set (with an id) or you can get all data sets.

### class TimeSeries:

**Description:** The class TimeSeries needs an id, a name, a description, the data and the legend. It also has functions to get several informations like the number of samples, the start date, the end date and the period.

### class LiveTimeSeries:

**Description:** The class LiveTimeSeries needs a name, description and a legend. It also has functions to get and to update the data.

### class TimeSeriesForecast:

**Description:** The class TimeSeriesForecast needs a name, a description, the training data, the forecasted data, the test data and the validation split.

## 4.4 FRONT END

The task of this module is to render the different html templates. The source code is located in the `frontend.py` file.

## 4.5 VISUALIZATION

The task of this module is to visualize the data. The source code is located in the `visualizations.py` file.

### **Time Series Plot:**

**Description:** This function generate the data to visualize a simple time series plot, with no additional given start and end date. We iterate through all data from the data set, the title will be set and we return a object with a title, the x accessor, the date, the y accessor, the value, the data, and the legend.

### **Add Rolling Mean:**

**Description:** This function generate the data to visualize a simple time series plot with an rolling mean. We calculate the rolling mean and generate the data object with the data and the legend.

### **Add Rolling Std:**

**Description:** This function generate the data to visualize a simple time series plot with an rolling standard deviation. We calculate the rolling standard deviation and generate the data object with the data and the legend.

### **Auto Correlation Plot:**

**Description:** This function generate the data to visualize a auto correlation time series plot with the `max_lag` parameter. We iterate through all data from the data set in range from 1 to the `max_lag` and generate the data object

### **Forecasting Eval Plot:**

**Description:** This function generate the data to visualize a forecasting time series plot.

### **Build Data Object:**

**Description:** This function build the object for all other functions in this class. The object consists of a date and a value.