# Science Project Title

Your name | Teacher's name | School

## Problem / Question

Enter your question here (statement of the problem)

## Hypothesis

- Add your answer / solution here
- Write hypothesis before you begin the experiment
- This should be your best educated guess based on your research

## Project Overview

- Add a brief overview or summary of your project. (Use the Bullets button on the Home tab to remove the bullets.)

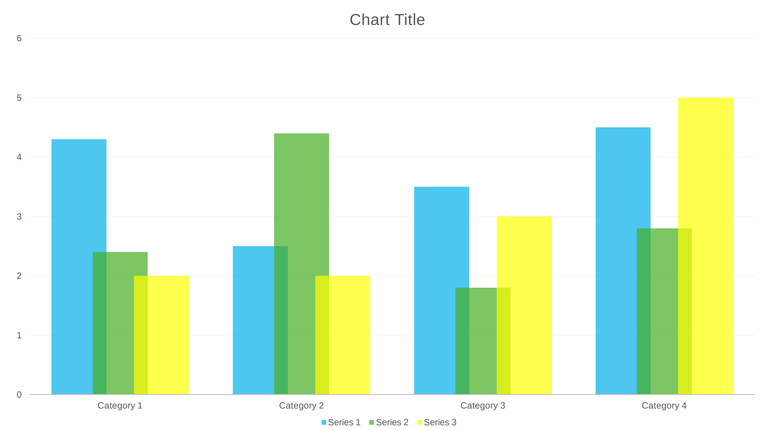
## Materials

| Materials (detailed list) | Quantity (be specific) |
|---|---|
| Item | Amount |
| Item | Amount |
| Item | Amount |
| Item | Amount |
| Item | Amount |
| Item | Amount |
| Item | Amount |

## Procedure

| Step 1 | Step 2 | Step 3 | Step 4 |
|---|---|---|---|
|  |  |  |  |
| Describe this step in your experiment | Describe this step in your experiment | Describe this step in your experiment | Describe this step in your experiment |

## Results




- Include results based on your experiments
- Result 2
- Result 3

## Conclusion

- Brief summary of what you discovered based on results
- Indicate and explain whether or not the data supports your hypothesis

## Variables / Research

| Controlled variables | Independent variable | Dependent variable |
|---|---|---|
| • These are kept the same throughout your experiments | • The **one** variable you purposely change and test | • The measure of change observed because of independent variable<br>• Decide how you will measure the change |

## Data / Observations

- Observation 1
- Observation 2
- Observation 3

## Works Cited

- Include print and electronic sources in alphabetical order

# Time Series Analysis and Forecasting

Dušan Stević SW10-2016

## Problem

Comparison between algorithms for time series forecasting. The main focus of the project is to study performance of statistical models (AR, MA, ARIMA) and machine learning models (LSTM).

## Motivation

- A time series is a series of data points indexed in time order.
- Time series forecasting is the use of a model to predict future values based on previously observed values.
- The motivation for this study is to present performance evaluation for state-of-the-art algorithms for time series forecasting.

## Objectives

- Data collection (Data acquisition)
- Data preprocessing
- Develop models
- Evaluate algorithmic performance

## Data collection and preprocessing

- Historical stock prices
- Stock market API
- Handling NA values
- Feature engineering and scaling

## Detecting stationarity in time series

### Rolling (moving) statistics

- Additive time series
- Multiplicative time series
- Visual Test
- Rolling Mean
- Rolling Standard Deviation

### Augmented Dickey Fuller Test

- Additive time series
- Statistical Tests
- p-value threshold 5%
- (H0): Time series is non-stationary. It has some time dependent structure.
- (H1): Time series is stationary

### Kwiatkowski–Phillips–Schmidt–Shin Test

- Multiplicative time series
- Statistical Tests
- p-value threshold 5%
- (H0): Time series is non-stationary. It has some time dependent structure.
- (H1): Time series is stationary

## Techniques for making time series stationary

- Differencing (Time Shift Transformation)
- Decomposition (Time series = Trend + Seasonality + Noise)

## Differencing

Given a set of observation on the time series:

$$x0, x1, x2, x3, \ldots . xn$$

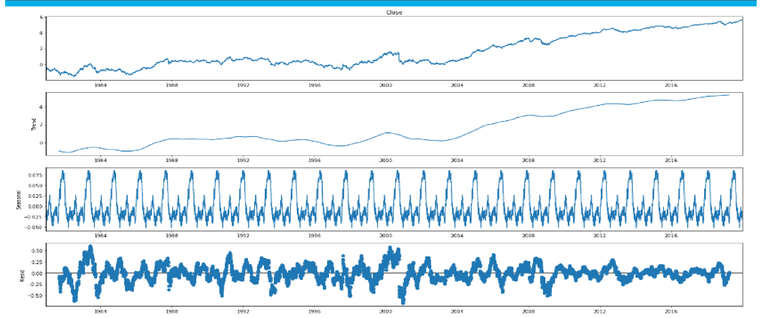
The shifted values will be:

$$null, x0, x1, x2, \ldots . xn \longleftarrow$$ - basically all xi's shifted by 1 pos to right

Thus, the time series with time shifted values are:

$$null, (x1 - x0), (x2 - x1), (x3 - x2), (x4 - x3), \ldots . (xn - x_{n-1})$$

## Decomposition




## Statistical Models

- The Autoregressive AR(p) Part of ARIMA : describes relationship between a current observation and previous observations.
- The Integrated I(d) Part of ARIMA: describes number of lags (shifts) until time series becomes stationary.
- The Moving Average MA(q) Part of ARIMA: describes relationship between an observation and a residual error from previous observations.