

# Statistics

**Spoken Tutorial Project**

**<http://spoken-tutorial.org>**

**National Mission on Education through ICT**

**<http://sakshat.ac.in>**

**Script:Thirumalesh H S**

**Narrator:Trupti Kini**

**IIT Bombay**

**18th Jan, 2018**



# Objectives



# Objectives

- ▶ **Do statistical operations in Python**



# Objectives

- ▶ Do statistical operations in Python
- ▶ **Sum a set of numbers**



# Objectives

- ▶ Do statistical operations in Python
- ▶ Sum a set of numbers
- ▶ Find their mean, median and standard deviation



# System Specifications

- ▶ **Ubuntu Linux 16.04 operating system**



# System Specifications

- ▶ **Ubuntu Linux 16.04** operating system
- ▶ **Python 3.4.3**



# System Specifications

- ▶ **Ubuntu Linux 16.04** operating system
- ▶ **Python 3.4.3**
- ▶ **IPython 5.1.0**



# Pre-requisites

**To practise this tutorial, you should know how to -**



# Pre-requisites

To practise this tutorial, you should know how to -

- ▶ Load data from files



# Pre-requisites

To practise this tutorial, you should know how to -

- ▶ Load data from files
- ▶ Use Lists



# Pre-requisites

To practise this tutorial, you should know how to -

- ▶ Load data from files
- ▶ Use Lists
- ▶ Access parts of Arrays



# Pre-requisites

To practise this tutorial, you should know how to -

- ▶ Load data from files
- ▶ Use Lists
- ▶ Access parts of Arrays

If not, see the pre-requisite Python tutorials on <http://spoken-tutorial.org>



# Numpy(Numerical Python)



# Numpy(Numerical Python)

- ▶ It is a library consisting of precompiled functions for mathematical and numerical routines



# Numpy(Numerical Python)

- ▶ It is a library consisting of precompiled functions for mathematical and numerical routines
- ▶ **NumPy has to be installed separately**



# Two-Dimensional array



# Two-Dimensional array

- ▶ **We can calculate the mean across each of the axis of the array**



# Two-Dimensional array

- ▶ We can calculate the mean across each of the axis of the array
- ▶ The axis of rows is referred by 0 and columns by 1



# Two-Dimensional array

- ▶ We can calculate the mean across each of the axis of the array
- ▶ The axis of rows is referred by **0** and columns by **1**
- ▶ To calculate mean across all columns, we have to pass extra parameter, **1** for the axis



# Exercise 1

- ▶ Refer to the file **football.txt**, that is available in the code files link of this tutorial
- ▶ Download and save the file in the present working directory



# Exercise 1

In `football.txt`, first column is **player name**, second is **goals at home** and third column is **goals away**

1. Find the total goals for each player
2. **Mean** of home and goals away
3. **Standard deviation** of home and goals away



# Summary

**We have learnt to do the standard statistical operations like:**

- ▶ **sum**
- ▶ **mean**
- ▶ **median**
- ▶ **standard deviation**



# Assignment

1. Given a two dimensional list as,  
`two_dimensional_list=[[3,5,8,2,1],[4,3,6,2,1]]`  
How do you calculate the **mean** of each row?
2. Calculate the **median** of the given list  
`student_marks=[74,78,56,87,91,82]`



# Assignment

3. There is a file with 6 columns.  
But we want to load text only  
from columns 2, 3, 4, 5.

How do we specify that?



# Solution

1. `np.mean(two_dimensional_list, 1)`
2. `np.median(student_marks)`
3. `usecols = (2, 3, 4, 5)`



# Forum to answer questions

- ▶ **Do you have questions in THIS Spoken Tutorial?**
- ▶ **Choose the minute and second where you have the question.**
- ▶ **Explain your question briefly.**
- ▶ **Someone from the FOSSEE team will answer them. Please visit**

<http://forums.spoken-tutorial.org/>



# Forum to answer questions

- ▶ Questions not related to the Spoken Tutorial?
- ▶ Do you have general / technical questions on the Software?
- ▶ Please visit the FOSSEE Forum  
<http://forums.fossee.in/>
- ▶ Choose the Software and post your question.



# Textbook Companion Project

- ▶ **The FOSSEE team coordinates coding of solved examples of popular books**
- ▶ **We give honorarium and certificate to those who do this**

**For more details, please visit this site:**

<http://tbc-python.fossee.in/>



# Acknowledgements

- ▶ **Spoken Tutorial Project is a part of the Talk to a Teacher project**
- ▶ **It is supported by the National Mission on Education through ICT, MHRD, Government of India**
- ▶ **More information on this Mission is available at:**

<http://spoken-tutorial.org/NMEICT-Intro>



# THANK YOU!

For more Information, visit our website  
<http://fossee.in/>

