

About Myself...



Name	Kathirmani Sukumar
Currently	Data Scientist @Gramener (4.5 years)
Previously	Project Associate @IIT Madras (3 years)
Total Experience	7.5 years
Areas of Expertise	Data Visualization, Data Science, Exploratory Data Analysis, Web development, Python, R
Also involved in	Corporate / Academic trainings , workshops, talks
Current location	Bangalore

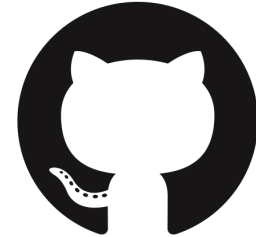
Social Media Profiles



<http://skathirmani.com>



<https://twitter.com/skathirmani>



<https://github.com/skathirmani>



<http://stackoverflow.com/users/1645853/kathirmani-sukumar>



<https://in.linkedin.com/in/skathirmani>

Talk at Amrita University on Exploratory Data Analysis



Talk at UpGrad on Data Science



Link: <http://blog.gramener.com/2514/introduction-to-exploratory-data-analysis-a-video-lesson>

Gramener blog
Data, Stories & Insight

SEARCH

Search ...

RECENT POSTS

[Introduction to Exploratory Data Analysis – A video lesson](#)

[Careers in Analytics](#)

[That's How We Sometimes Surprise Our New Joiners, Welcome Party !](#)

[Dark Patterns in Web Design](#)

[Gramener at CapGemini Startup Day](#)

[Gramener Boot Camp and Outing](#)

Introduction to Exploratory Data Analysis – A video lesson

It's always not necessary to use machine learning algorithm for extracting interesting stories from data. Here is a video by Kathirmani Sukumar, Data Scientist at Gramener, which explains data analysis sans complicated machine learning techniques. The video lesson will help in learning how to analyse data using a few simple (but powerful) techniques based on Exploratory Data Analysis (EDA) using data from the sport of Cricket. It is useful for those who want to do data analysis, but are not sure of where to start and what to learn from. It also discusses about basics of data types, data mutation and univariate analysis, and these techniques are domain agnostic. One can apply the same techniques on any data from any domain. It focuses on using Pandas library for data processing and plotting the results. The Jupyter notebook can be downloaded from <http://bit.ly/2hCJrqY>. The next video in the series will be a lesson on univariate and bivariate analysis.

Cricket Analysis

localhost:8888/notebooks/Cricket%20Analysis.ipynb#Average-runs-by-top-players

jupyter Cricket Analysis Last Checkpoint: 10/23/2016 (autosaved)

File Edit View Insert Cell Kernel Help Python [Root]

Exploratory Data Analysis (EDA)

In [2]:

```
import pandas as pd
import matplotlib
import matplotlib.pyplot as plt
matplotlib inline
import numpy as np
import seaborn
matplotlib.style.use('ggplot')
from scipy.stats import ttest_ind
from scipy.stats import chi2_contingency
from sklearn.cluster import KMeans
```

Look at your data
Understand what is each row and column is all about?

Common data types

- Dimensions - Columns on which you can group by (Country, Departments, Age group, College name etc.)
- Metrics - All numeric columns (Revenue, Marks, Volume etc.)
- Dates - (Date of birth, Purchased date etc.)
- Text - (Product description, user reviews, tweets etc.)
- Geo - (Lat long, Country, District, pin codes etc)

Awards for the product that I am currently leading at Gramener



Received award from India's commerce minister Nirmal Sitharaman during Express IT Awards organized by Financial Express for the product that I am currently leading.



Received "Best Practices in Data Analysis & Strategy" in Big Data Analytics & Insights Summit



Contents lists available at SciVerse ScienceDirect

NDT&E International

journal homepage: www.elsevier.com/locate/ndteint



Online data compression of MFL signals for pipeline inspection

S. Kathirmani^a, A.K. Tangirala^{a,*}, S. Saha^b, S. Mukhopadhyay^b

^a Department of Chemical Engineering, Indian Institute of Technology Madras, India

^b Control Instrumentation Division, Bhabha Atomic Research Centre, Mumbai, India

ARTICLE INFO

Article history:

Received 30 August 2011

Received in revised form

25 September 2011

ABSTRACT

The paper presents a novel three-stage algorithm for online compression of magnetic flux leakage (MFL) signals that are acquired in inspection of oil and gas pipelines. In the first stage, blocks of MFL signal are screened for useful information using a semi-robust statistical measure, Mean Absolute