# My Project

Generated by Doxygen 1.8.13

# Contents

# Chapter 1

# Namespace Index

## 1.1 Namespace List

Here is a list of all documented namespaces with brief descriptions:

# Chapter 2

# Namespace Documentation

## 2.1 title_eng Namespace Reference

### Variables

- bool firstline = True

  *Skip the first line, as it is a header.*
- title_file = open('dataset_with_titleEng.csv','a+')

  *dataset extended with titles(in English)*
- string **comma** = ','
- string **sep** = '('
- int url_column = 4

  *url column in kaggle dataset*
- readCSV = csv.reader(csvfile, delimiter=',')

  *Initial dataset taken from Kaggle.*
- res = requests.get(row[url_column])

  *fetch the url of Kaggle dataset, one at a time*
- soup = bs4.BeautifulSoup(res.text,'lxml')

  *check if response object is not null*
- **title_text** = soup.select('title')
- **title** = title_text[0].getText()
- string old_entry = ','.join(map(str,row))

  *write old row entry, then add title column*

### 2.1.1 Detailed Description

```
This module takes input the Kaggle dataset, named output.csv and gives
the column extended dataset with English title added.
For web scraping : BeautifulSoup was used
```

### 2.1.2 Variable Documentation

#### 2.1.2.1 soup

```
title_eng.soup = bs4.BeautifulSoup(res.text,'lxml')
```

check if response object is not null

use bs4 to

# Index