

Plot the Graph for the total number of issues with different labels created and closed on every date on GitHub for SPM587SP21 issues project

Example of Issue Form Filled out

Author: SCM587SP21 created_at: 2021-03-03, labels: Category:Bug, DetectionPhase:Field, OriginationPhase:Coding, Priority:Critical State: closed, issue_number: 002, closed_at: 2021-03-05

Data Set File: SPM587SP21issues.json

```
In [1]: import os

import warnings
warnings.filterwarnings('ignore')

import pandas as pd # panda's nickname is pd

import numpy as np # numpy as np

from pandas import DataFrame, Series # for convenience

import matplotlib.pyplot as plt

%matplotlib inline
```

Requirement : Get the total number of issues with different labels for every date and plot them in a stacked chart

```
In [2]: # Read the JSON file into a list of dictionaries

import json
list_of_issues_dict_data = [json.loads(line) for line in open('SPM587SP21issues.json')]
```

```
In [3]: # Create the DataFrame object for the list_of_issues_dict_data object

issues_df = DataFrame(list_of_issues_dict_data)
```

```
In [4]: # Sanity test: print rows in our DataFrame

issues_df
```

Out[4]:

	issue_number	created_at	closed_at	labels	State	Author
0	322	2021-04-07	2021-04-08	[Category:Enhancement, DetectionPhase:Testing,...	closed	MSP21SCM41J
1	320	2021-04-07	2021-04-08	[Category:Inquiry, DetectionPhase:Field, Origina...	closed	SSP21SCM84S
2	317	2021-04-07	2021-04-07	[Category:Bug, DetectionPhase:Testing, Origina...	closed	RSP21SCM47M
3	314	2021-04-07	2021-04-07	[Category:Enhancement, DetectionPhase:Testing,...	closed	RSP21SCM43S
4	312	2021-04-07	2021-04-07	[Category:Inquiry, DetectionPhase:Field, Origina...	closed	YSP21SCM14W
...
155	5	2021-03-03	None	[Category:Enhancement, DetectionPhase:Requirem...	open	SPM587SP21
156	4	2021-03-03	None	[Category:Bug, DetectionPhase:Field, Originati...	open	SPM587SP21
157	3	2021-03-03	2021-03-05	[Category:Inquiry, DetectionPhase:Documentatio...	closed	SPM587SP21
158	2	2021-03-03	2021-03-05	[Category:Enhancement, DetectionPhase:Design, ...	closed	SPM587SP21
159	1	2021-03-03	None	[Category:Bug, DetectionPhase:Coding, Originat...	open	SPM587SP21

160 rows × 6 columns

```
In [5]: # Prepare and Clean the dataframe object

wrangled_issues_df = issues_df[['Author', 'State', 'closed_at', 'created_at', 'issue_number', 'labels']]
wrangled_issues_df.loc[0:len(wrangled_issues_df), 'OriginationPhase'] = np.NaN
wrangled_issues_df.loc[0:len(wrangled_issues_df), 'DetectionPhase'] = np.NaN
wrangled_issues_df.loc[0:len(wrangled_issues_df), 'Category'] = np.NaN
wrangled_issues_df.loc[0:len(wrangled_issues_df), 'Priority'] = np.NaN
wrangled_issues_df.loc[0:len(wrangled_issues_df), 'Status'] = np.NaN
```

```
In [6]: wrangled_issues_df
```

Out[6]:

	Author	State	closed_at	created_at	issue_number	labels	OriginationPhase	DetectionPhase	Catego
0	MSP21SCM41J	closed	2021-04-08	2021-04-07	322	[Category:Enhancement, DetectionPhase:Testing,...	NaN	NaN	N
1	SSP21SCM84S	closed	2021-04-08	2021-04-07	320	[Category:Inquiry, DetectionPhase:Field, Origina...	NaN	NaN	N
2	RSP21SCM47M	closed	2021-04-07	2021-04-07	317	[Category:Bug, DetectionPhase:Testing, Origina...	NaN	NaN	N
3	RSP21SCM43S	closed	2021-04-07	2021-04-07	314	[Category:Enhancement, DetectionPhase:Testing,...	NaN	NaN	N
4	YSP21SCM14W	closed	2021-04-07	2021-04-07	312	[Category:Inquiry, DetectionPhase:Field, Origina...	NaN	NaN	N
...
155	SPM587SP21	open	None	2021-03-03	5	[Category:Enhancement, DetectionPhase:Requirem...	NaN	NaN	N
156	SPM587SP21	open	None	2021-03-03	4	[Category:Bug, DetectionPhase:Field, Originati...	NaN	NaN	N
157	SPM587SP21	closed	2021-03-05	2021-03-03	3	[Category:Inquiry, DetectionPhase:Documentatio...	NaN	NaN	N
158	SPM587SP21	closed	2021-03-05	2021-03-03	2	[Category:Enhancement, DetectionPhase:Design, ...	NaN	NaN	N
159	SPM587SP21	open	None	2021-03-03	1	[Category:Bug, DetectionPhase:Coding, Originat...	NaN	NaN	N

160 rows × 11 columns

```
In [7]: for i in range(0, len(wrangled_issues_df)):
        if wrangled_issues_df.iloc[i]['labels']:
            for label in wrangled_issues_df.iloc[i]['labels']:
                if len(label.split(':')) == 2:
                    label_name= (label.split(':')[0])
                    label_value= (label.split(':')[1])
                    wrangled_issues_df.loc[i, label_name]=label_value
```

```
In [8]: wrangled_issues_df
```

Out[8]:

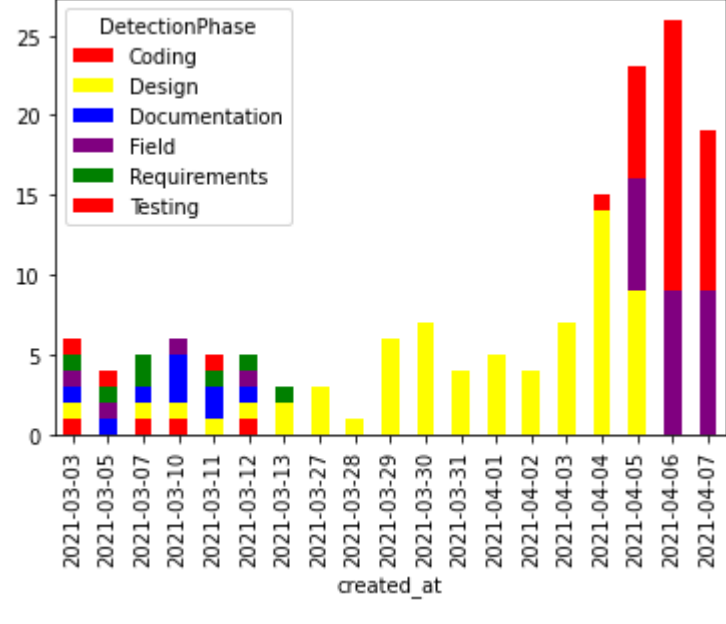
	Author	State	closed_at	created_at	issue_number	labels	OriginationPhase	DetectionPhase	Cs
0	MSP21SCM41J	closed	2021-04-08	2021-04-07	322	[Category:Enhancement, DetectionPhase:Testing,...	Requirements	Testing	Enhan
1	SSP21SCM84S	closed	2021-04-08	2021-04-07	320	[Category:Inquiry, DetectionPhase:Field, Origina...	Coding	Field	
2	RSP21SCM47M	closed	2021-04-07	2021-04-07	317	[Category:Bug, DetectionPhase:Testing, Origina...	Requirements	Testing	
3	RSP21SCM43S	closed	2021-04-07	2021-04-07	314	[Category:Enhancement, DetectionPhase:Testing,...	Design	Testing	Enhan
4	YSP21SCM14W	closed	2021-04-07	2021-04-07	312	[Category:Inquiry, DetectionPhase:Field, Origina...	Documentation	Field	
...
155	SPM587SP21	open	None	2021-03-03	5	[Category:Enhancement, DetectionPhase:Requirem...	Requirements	Requirements	Enhan
156	SPM587SP21	open	None	2021-03-03	4	[Category:Bug, DetectionPhase:Field, Originati...	Field	Field	
157	SPM587SP21	closed	2021-03-05	2021-03-03	3	[Category:Inquiry, DetectionPhase:Documentatio...	Documentation	Documentation	
158	SPM587SP21	closed	2021-03-05	2021-03-03	2	[Category:Enhancement, DetectionPhase:Design, ...	Design	Design	Enhan
159	SPM587SP21	open	None	2021-03-03	1	[Category:Bug, DetectionPhase:Coding, Originat...	Coding	Coding	

160 rows × 11 columns

```
In [9]: # Plot in Bar Chart the total number of issues created every day for every Detaction Phase

LabelsReviewedByDate = wrangled_issues_df.groupby(['created_at', 'DetectionPhase']).created_at.count()

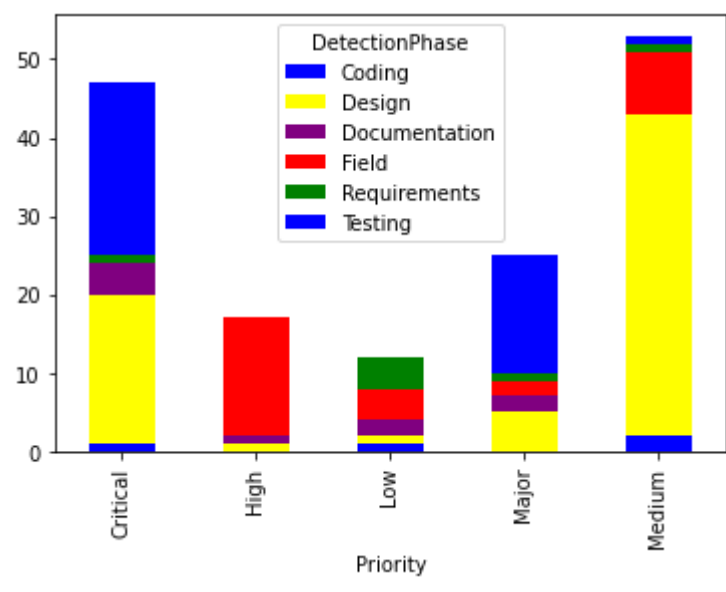
dateLabelsFig = LabelsReviewedByDate.unstack().plot(kind='bar', stacked=True, color=['red', 'yellow', 'blue', 'purple', 'green'], grid=False)
```



```
In [10]: # Plot in Bar Chart the total number of issues created every Phase based on their priorities

LabelsReviewedByDate = wrangled_issues_df.groupby(['Priority', 'DetectionPhase']).created_at.count()

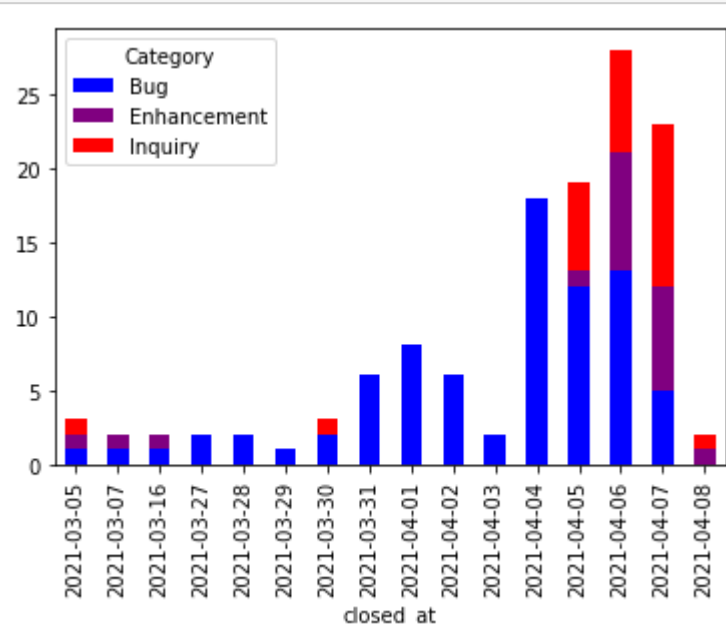
dateLabelsFig = LabelsReviewedByDate.unstack().plot(kind='bar', stacked=True, color=['blue', 'yellow', 'purple', 'red', 'green'], grid=False)
```



```
In [11]: # Plot in Bar Chart the total number of issues closed every day for every Category

LabelsReviewedByDate = wrangled_issues_df.groupby(['closed_at', 'Category']).closed_at.count()

dateLabelsFig = LabelsReviewedByDate.unstack().plot(kind='bar', stacked=True, color=['blue', 'purple', 'red', 'yellow', 'green'], grid=False)
```

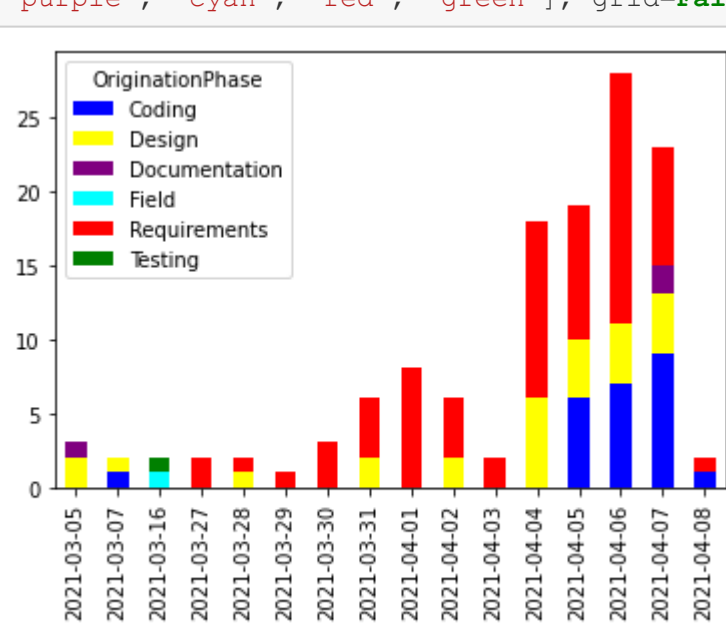


Requirement #1: Plot in Bar Chart the total number of issues closed every day for every Origination Phase

```
In [14]: # Requirement #1: Add your code here

LabelsReviewedByDate = wrangled_issues_df.groupby(['closed_at', 'OriginationPhase']).closed_at.count()

dateLabelsFig = LabelsReviewedByDate.unstack().plot(kind='bar', stacked=True, color=['blue', 'yellow', 'purple', 'cyan', 'red', 'green'], grid=False)
```

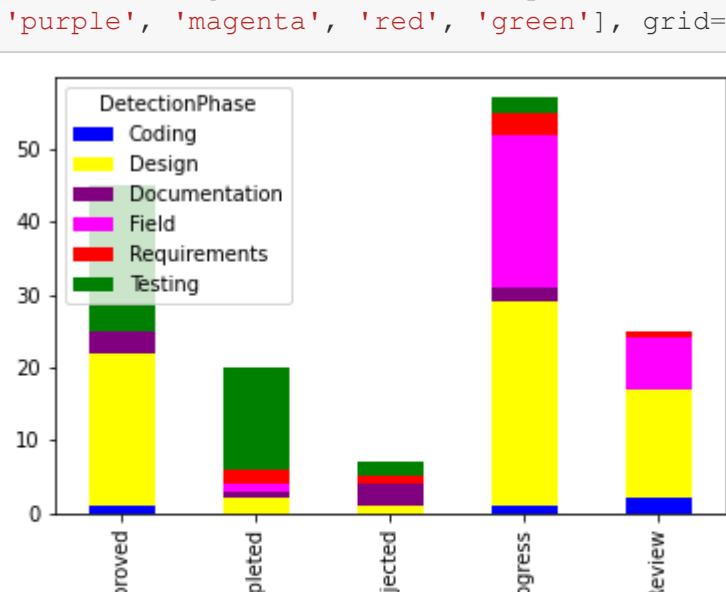


Requirement #2: # Plot in Bar Chart the total number of issues created for every Phase based on their Status

```
In [15]: # Requirement #2: Add your code here

LabelsReviewedByDate = wrangled_issues_df.groupby(['Status', 'DetectionPhase']).created_at.count()

dateLabelsFig = LabelsReviewedByDate.unstack().plot(kind='bar', stacked=True, color=['blue', 'yellow', 'purple', 'magenta', 'red', 'green'], grid=False)
```



```
In [ ]:
```