

Plot the Graph for the total number of issues with different labels created and closed on every date on GitHub for SPM587SP23 issues project

Example of Issue Form Filled out

Author: SPM587SP23 created_at: 2023-03-23, labels: Category:Enhancement, OriginationPhase:Requirements, Priority:Critical State: open, issue_number: 6, closed_at: 2023-03-23

Data Set File: SPM587FA22issues.json

```
In [15]: 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 [16]: # Read the JSON file into a list of dictionaries

import json
list_of_issues_dict_data = [json.loads(line) for line in open('SPM587SP23issues.json')]

In [17]: # Create the DataFrame object for the list_of_issues_dict_data object

issues_df = DataFrame(list_of_issues_dict_data)

In [18]: # Sanity test: print rows in our DataFrame

issues_df
```

Out[18]:

	issue_number	created_at	closed_at	labels	State	Author
0	497	2023-03-31	None	[Category: Inquiry, DetectionPhase: Field, Ori...	open	HSP23SCM78S
1	496	2023-03-31	None	[Category: Enhancement, DetectionPhase: Testin...	open	YSP23SCM71W
2	495	2023-03-30	None	[Category: Bug, DetectionPhase: Design, Origin...	open	PSP23SCM34E
3	493	2023-03-30	None	[Category: Enhancement, DetectionPhase: Testin...	open	SSP23SCM29R
4	489	2023-03-30	None	[Category: Bug, DetectionPhase: Design, Origin...	open	HSP23SCM78S
...
247	6	2023-03-23	2023-03-26	[Category: Enhancement, OriginationPhase: Requ...	closed	SPM587SP23
248	5	2023-03-23	2023-03-26	[DetectionPhase: Design, OriginationPhase: Des...	closed	SPM587SP23
249	4	2023-03-23	2023-03-27	[DetectionPhase: Field, DetectionPhase: Testin...	closed	SPM587SP23
250	3	2023-03-05	2023-03-26	[]	closed	SPM587SP23
251	1	2023-03-03	2023-03-23	[Category: Bug, DetectionPhase: Coding]	closed	SPM587SP23

252 rows × 6 columns

In [19]: *# Prepare and Clean the dataframe object*

```

wrangled_issues_df = issues_df[['Author', 'State', 'closed_at', 'created_at', 'issue_nu
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 [20]: wrangled_issues_df

Out[20]:

	Author	State	closed_at	created_at	issue_number	labels	OriginationPhase
0	HSP23SCM78S	open	None	2023-03-31	497	[Category: Inquiry, DetectionPhase: Field, Ori...	Na
1	YSP23SCM71W	open	None	2023-03-31	496	[Category: Enhancement, DetectionPhase: Testin...	Na
2	PSP23SCM34E	open	None	2023-03-30	495	[Category: Bug, DetectionPhase: Design, Origin...	Na
3	SSP23SCM29R	open	None	2023-03-30	493	[Category: Enhancement, DetectionPhase: Testin...	Na
4	HSP23SCM78S	open	None	2023-03-30	489	[Category: Bug, DetectionPhase: Design, Origin...	Na
...
247	SPM587SP23	closed	2023-03-26	2023-03-23	6	[Category: Enhancement, OriginationPhase: Requ...	Na
248	SPM587SP23	closed	2023-03-26	2023-03-23	5	[DetectionPhase: Design, OriginationPhase: Des...	Na
249	SPM587SP23	closed	2023-03-27	2023-03-23	4	[DetectionPhase: Field, DetectionPhase: Testin...	Na
250	SPM587SP23	closed	2023-03-26	2023-03-05	3	[]	Na
251	SPM587SP23	closed	2023-03-23	2023-03-03	1	[Category: Bug, DetectionPhase: Coding]	Na

252 rows × 11 columns



```
In [21]: 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 [22]: wrangled_issues_df
```

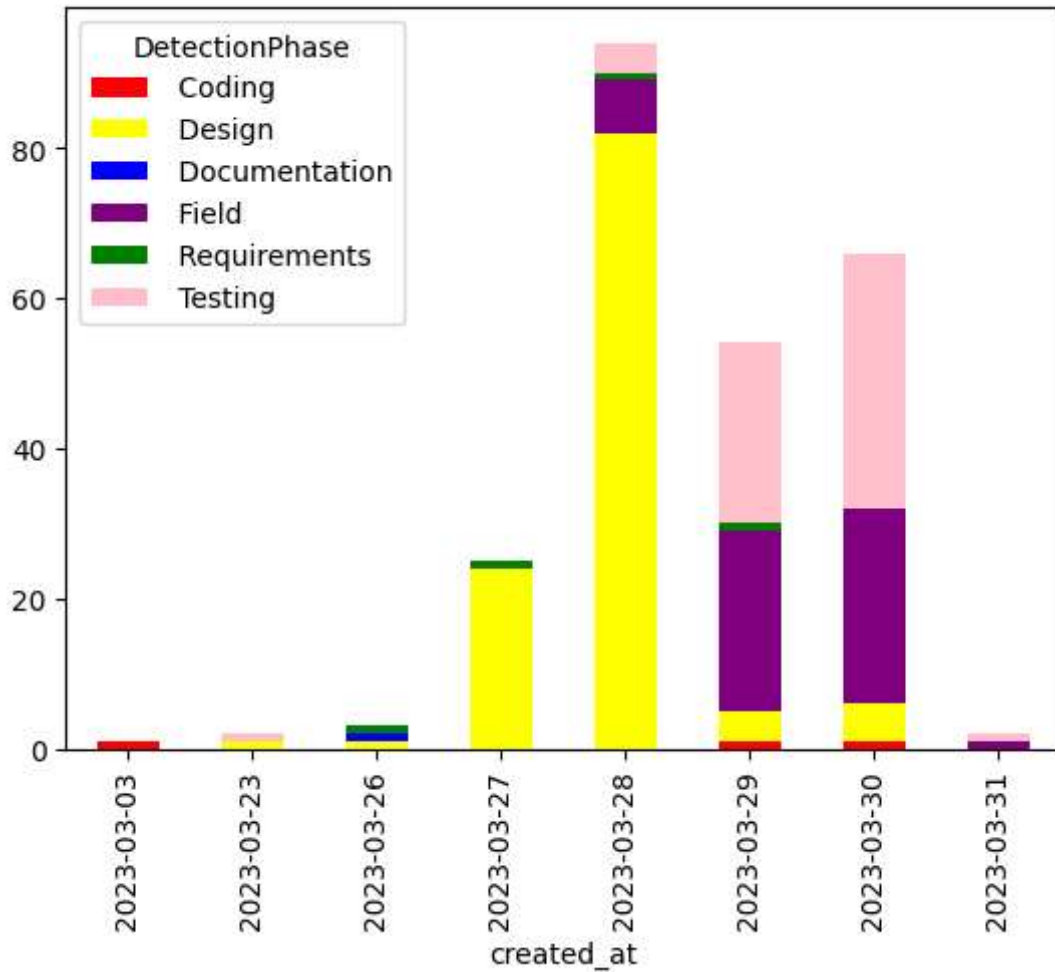
Out[22]:

	Author	State	closed_at	created_at	issue_number	labels	OriginationPhase
0	HSP23SCM78S	open	None	2023-03-31	497	[Category: Inquiry, DetectionPhase: Field, Ori...	Coding
1	YSP23SCM71W	open	None	2023-03-31	496	[Category: Enhancement, DetectionPhase: Testin...	Requirement
2	PSP23SCM34E	open	None	2023-03-30	495	[Category: Bug, DetectionPhase: Design, Origin...	Documentation
3	SSP23SCM29R	open	None	2023-03-30	493	[Category: Enhancement, DetectionPhase: Testin...	Design
4	HSP23SCM78S	open	None	2023-03-30	489	[Category: Bug, DetectionPhase: Design, Origin...	Design
...
247	SPM587SP23	closed	2023-03-26	2023-03-23	6	[Category: Enhancement, OriginationPhase: Requ...	Requirement
248	SPM587SP23	closed	2023-03-26	2023-03-23	5	[DetectionPhase: Design, OriginationPhase: Des...	Design
249	SPM587SP23	closed	2023-03-27	2023-03-23	4	[DetectionPhase: Field, DetectionPhase: Testin...	Testin
250	SPM587SP23	closed	2023-03-26	2023-03-05	3	[]	Na
251	SPM587SP23	closed	2023-03-23	2023-03-03	1	[Category: Bug, DetectionPhase: Coding]	Na

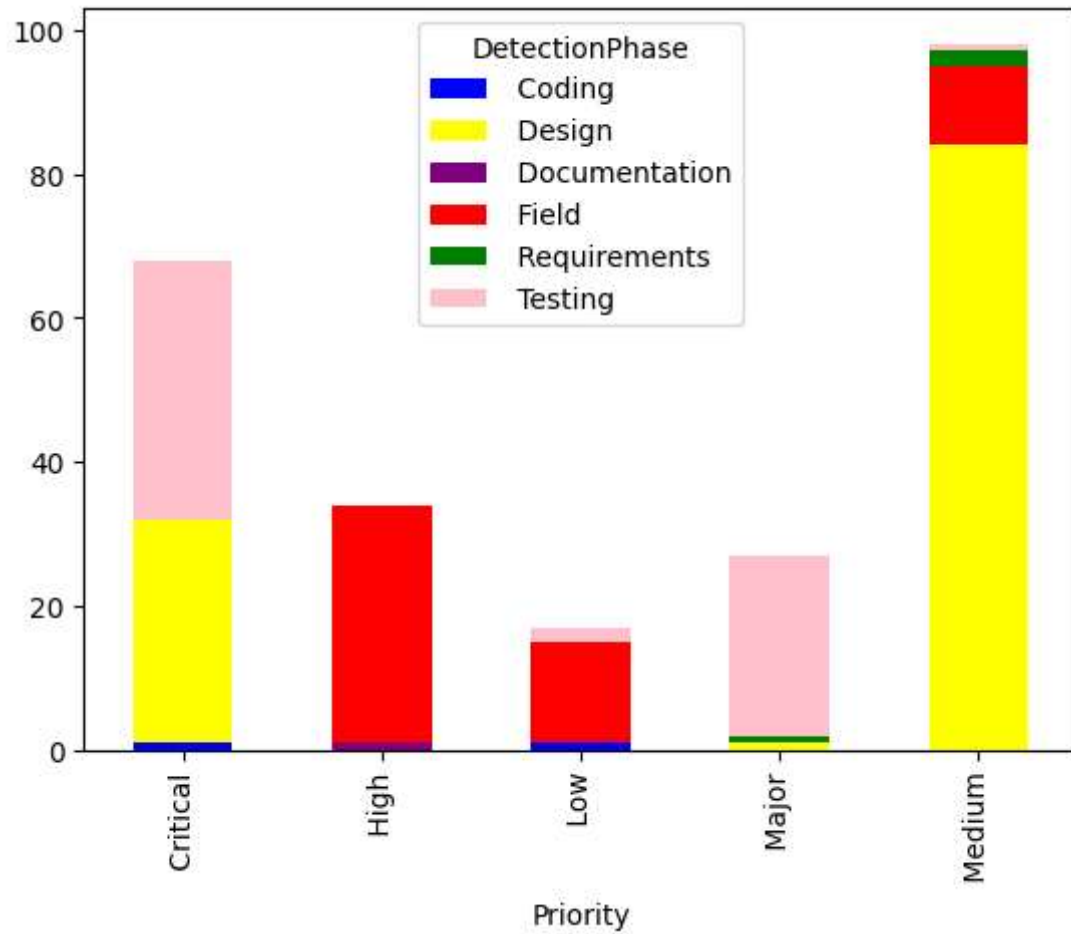
252 rows × 11 columns

```
In [23]: # Plot in Bar Chart the total number of issues created every day for every DetectionPhase
LabelsReviewedByDate = wrangled_issues_df.groupby(['created_at', 'DetectionPhase']).
```

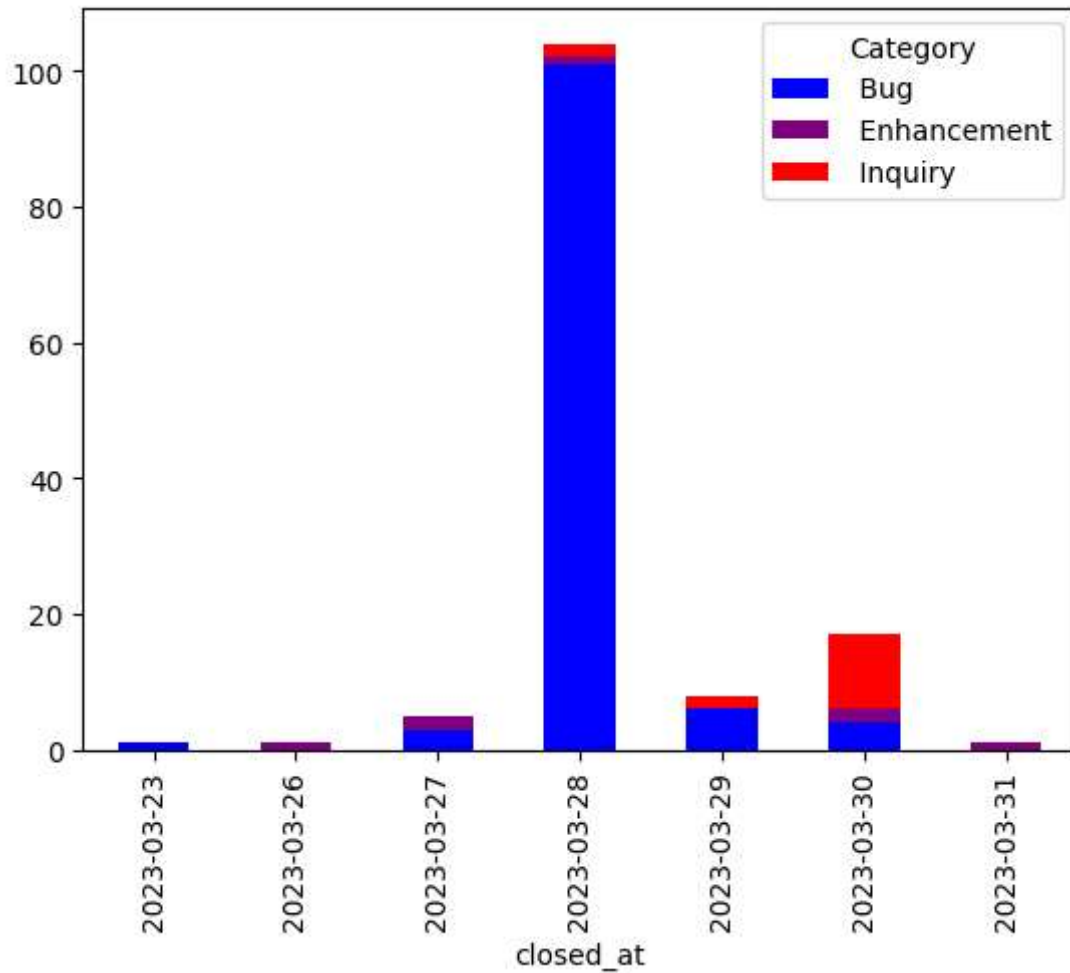
```
dateLabelsFig = LabelsReviewedByDate.unstack().plot(kind='bar',stacked=True, color
```



```
In [24]: # Plot in Bar Chart the total number of issues created for every Phase based on the
LabelsReviewedByDate = wrangled_issues_df.groupby(['Priority','DetectionPhase']).cr
dateLabelsFig = LabelsReviewedByDate.unstack().plot(kind='bar',stacked=True, color
```



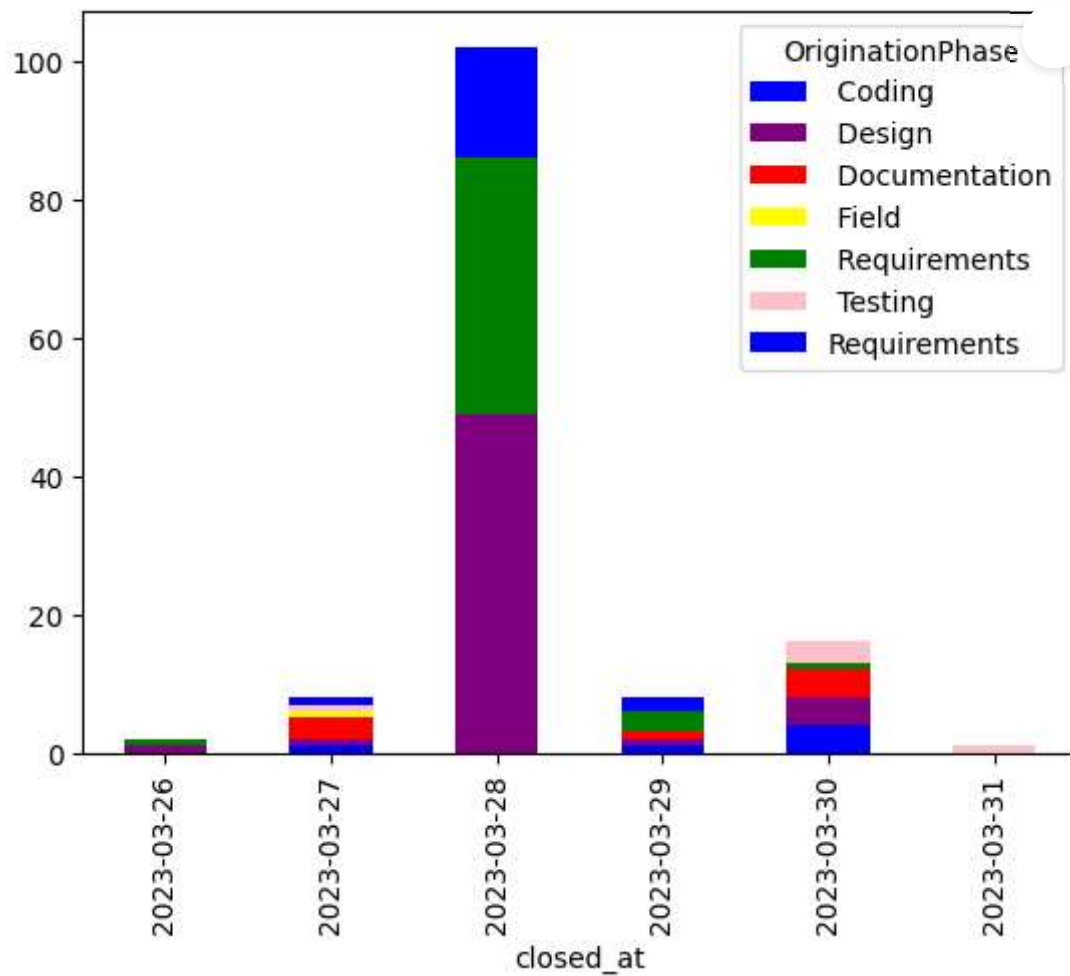
```
In [25]: # Plot in Bar Chart the total number of issues closed every day for every Category
LabelsReviewedByDate = wrangled_issues_df.groupby(['closed_at', 'Category']).closed_
dateLabelsFig = LabelsReviewedByDate.unstack().plot(kind='bar', stacked=True, color
```



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

```
In [26]: # Requirement #1: Add your code here

# Plot in Bar Chart the total number of issues closed every day for every Origination Phase
LabelsReviewedByDate = wrangled_issues_df.groupby(['closed_at', 'OriginationPhase'])
dateLabelsFig = LabelsReviewedByDate.unstack().plot(kind='bar', stacked=True, color
```

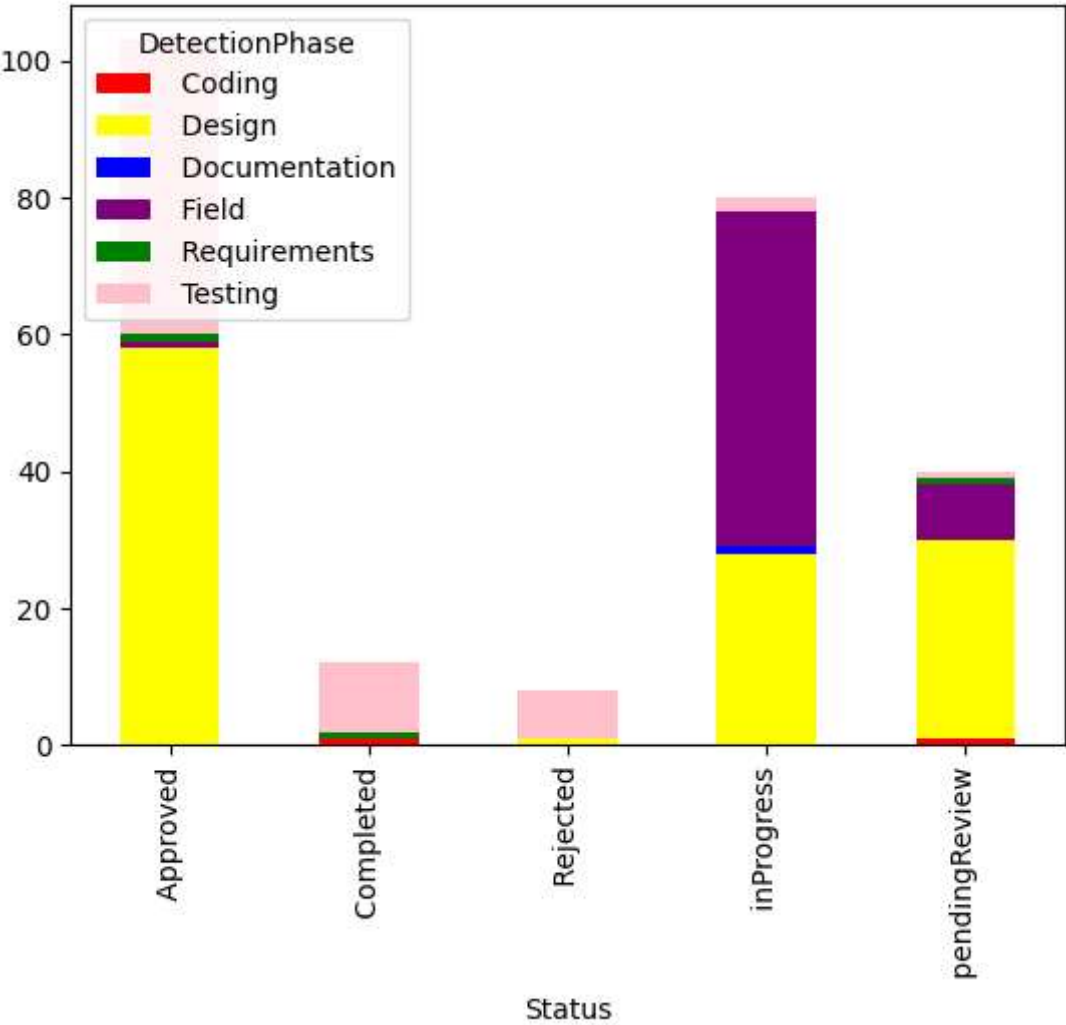


Requirement #2: # Plot in Bar Chart the total number of issues created for every Phase based on their Status

```
In [27]: # Requirement #2: Add your code here

# Plot in Bar Chart the total number of issues created every day for every Status

LabelsReviewedByDate = wrangled_issues_df.groupby(['Status', 'DetectionPhase']).crea
dateLabelsFig = LabelsReviewedByDate.unstack().plot(kind='bar', stacked=True, color
```

```
In [ ]:
In [ ]:
In [ ]:
In [ ]:
In [ ]:
In [ ]:
In [ ]:
In [ ]:
```