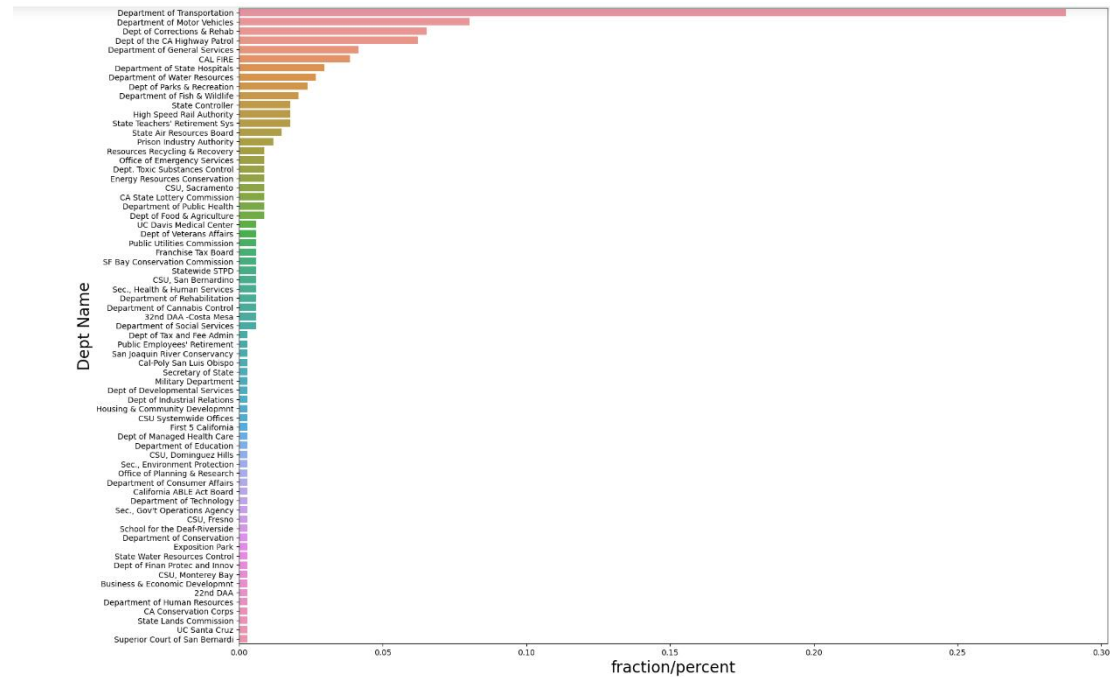


# California State Contracts Register

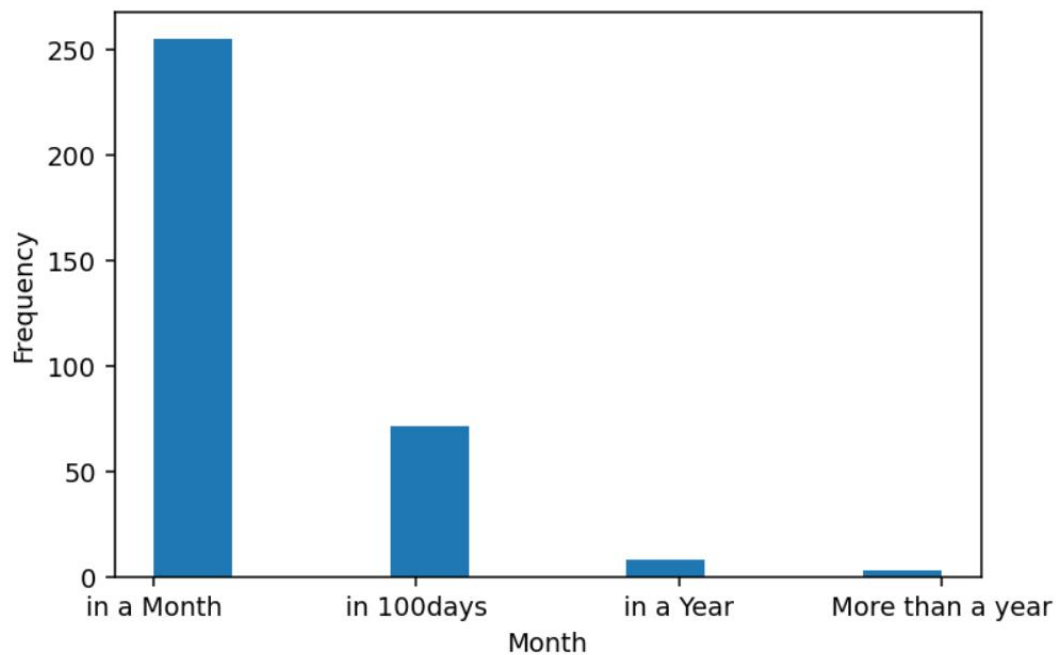
## Summary:

## Non-technical:

1. There are 337 contracts found in the register
2. It looks like department of Transportation has posted more contracts than others. It is way ahead of its peers.

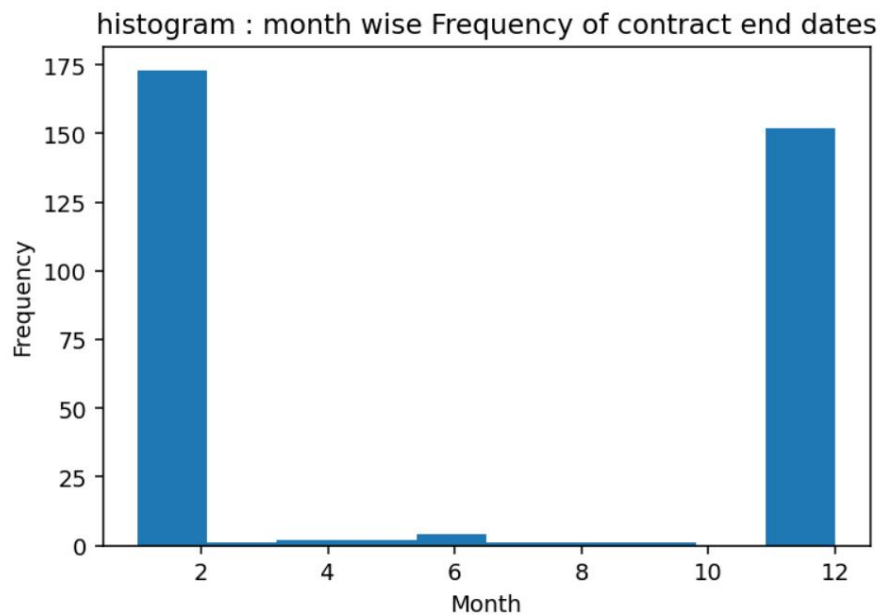


3. Most of the projects end in the next 30 days and very few projects have time of an year and above.



# California State Contracts Register

4. If we analyze why there are more projects that are going to end in a month, it looks like they mostly follow calendar year cycle for their projects i.e they close the contracts by dec/jan and give fresh orders there after. This is just an intelligent guess based on the data at hand and we can confirm this only if we have start dates of contracts.



Technical details:

There are two ways of getting data for this particular website:

1. Scrapping the table provided in the web page.
  - a) Class: DataDownloader()
  - b) File: DE\_WS\_task\_w1.py
  - c) Jupyter file: DE\_WS\_task\_w1.ipynb
  - d) Data: file1.csv
2. Direct download of the excel file provided in the same page
  - a) class DataDownloader2()
  - b) File: DE\_WS\_task\_w2.py
  - c) Jupyter file: DE\_WS\_task\_w2. ipynb
  - d) Data: ps.xls

How to use the code:

Create an instance of class and run the module by passing url and webdriver path(raw format-> r'path')