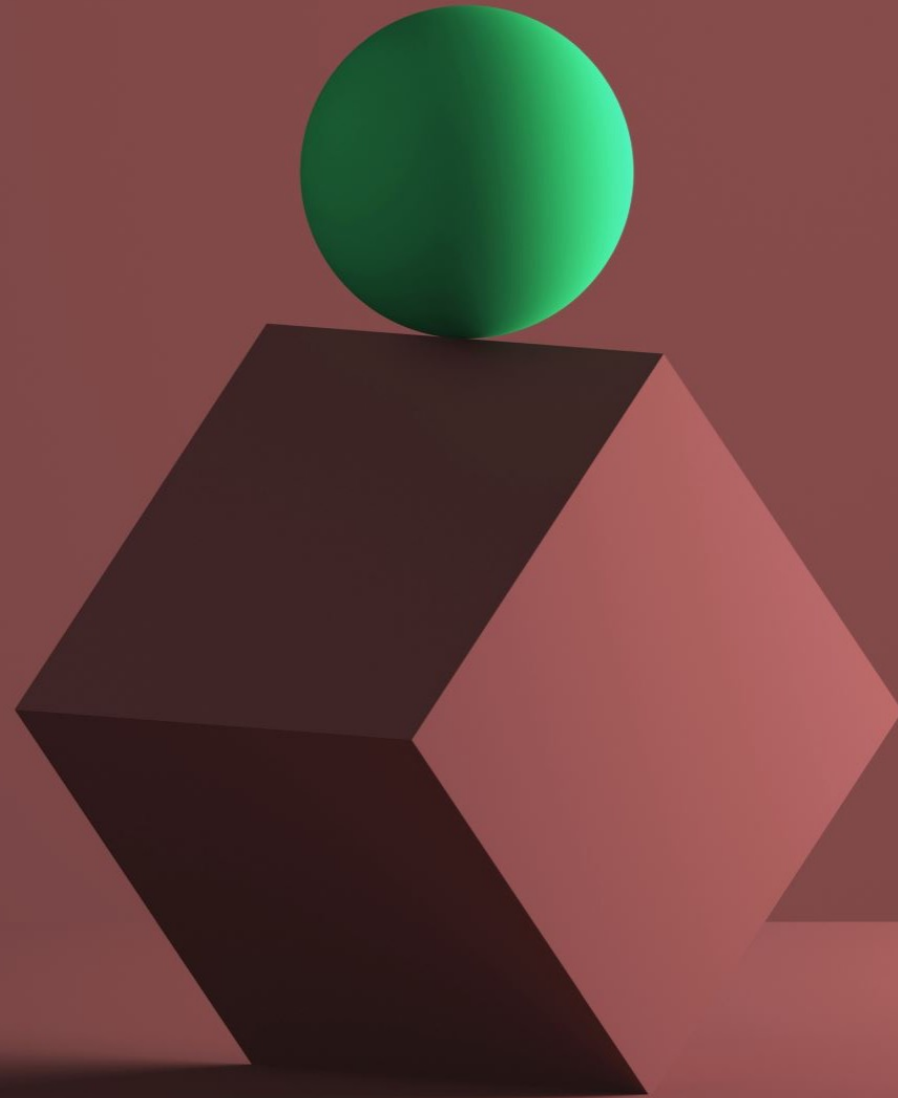# A Presentation On Data Analysis

A Final Project on Descriptive Statistics and Visualization.

- Introduction to Data Analytics (ALY6000)

- **Guided by:**
  Prof. Mohammad Shafiqul Islam

- **Submitted by**:
  Name : Pratikkumar Indravadan Malaviya
  NUID : 002963548

- **Date of submission :**
  20th February'2022

N Northeastern University

AGENDA

DATASET
- Information of Dataset
- Data Understanding

DATA ANALYSIS
- Descriptive Statistics
- Plotted Graphs

SUMMARY
- Discussion

DATASET

Cyber Security Breach

- The collected data set is belonging to the field of Cyber Security where, several Data Breaches happened in USA.

Structure of Dataset

- The whole data encapsulated in .csv file.
- It contains total ~**1054 observations** which are divided into **10 variables**.

DATA UNDERSTANDING

> Cyber Security Breaches are increasing in USA from 2009 to 2014.

> Apart from all states, major ones were the active victims of cyber breach.

> The state of **California** has accounted highest cyber breaches.

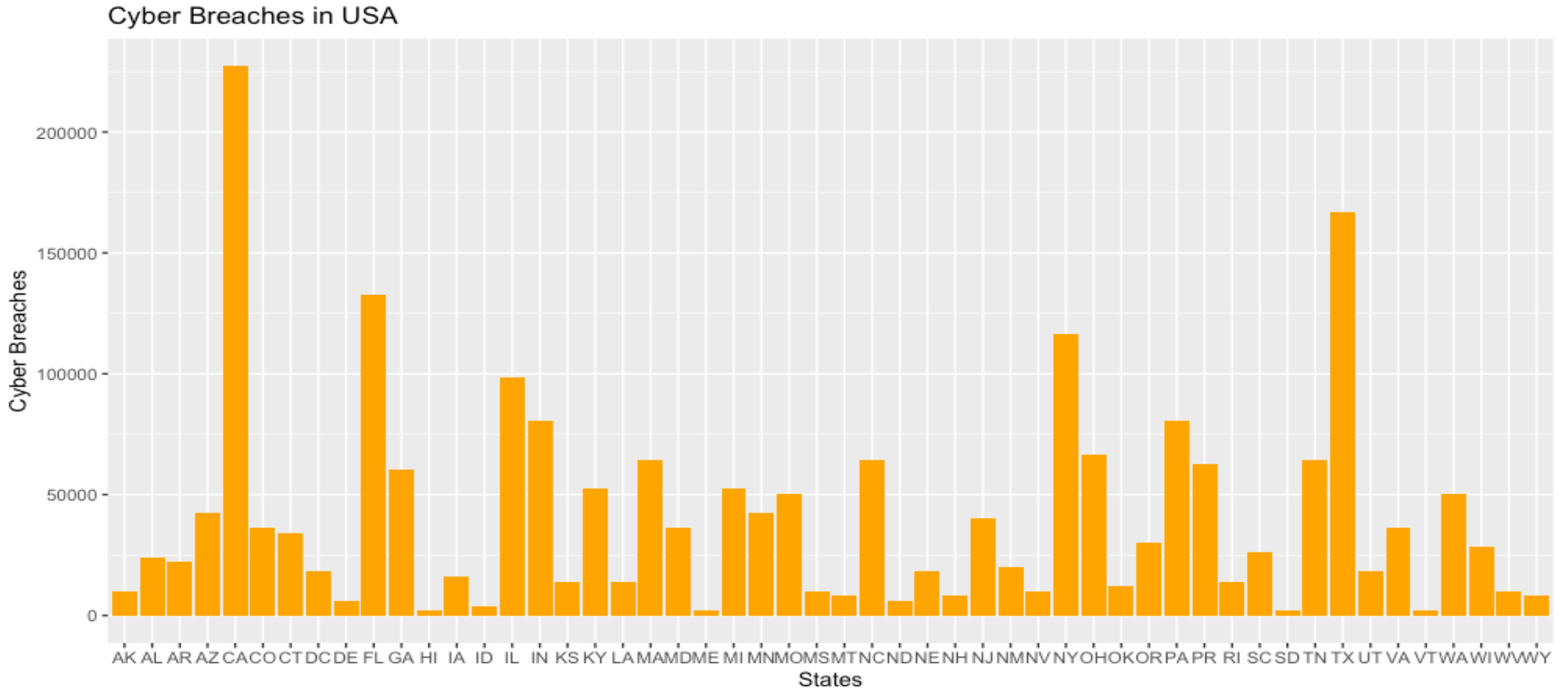> Several business associations were involved who passed the data. *(Mediators)*

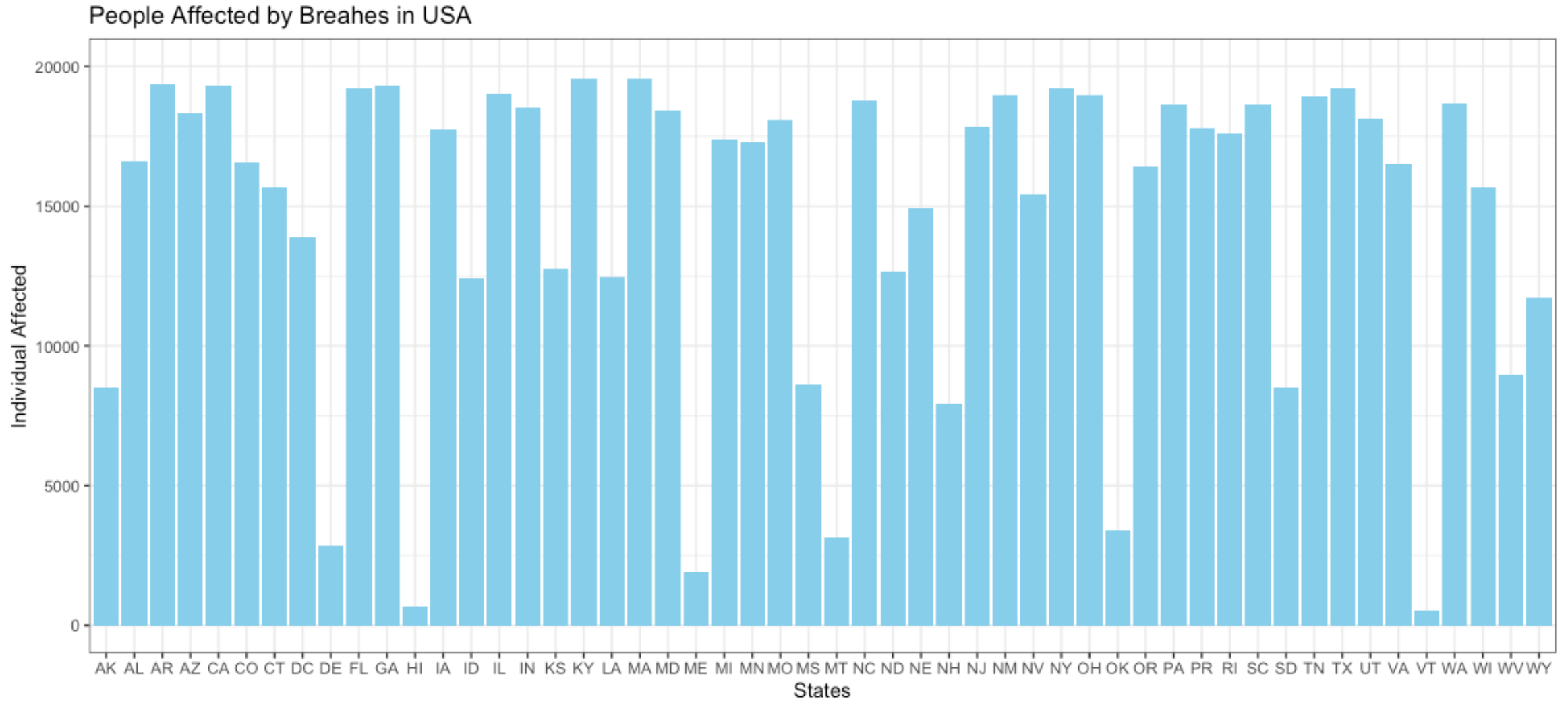> Brief information about types of breach and located devices.

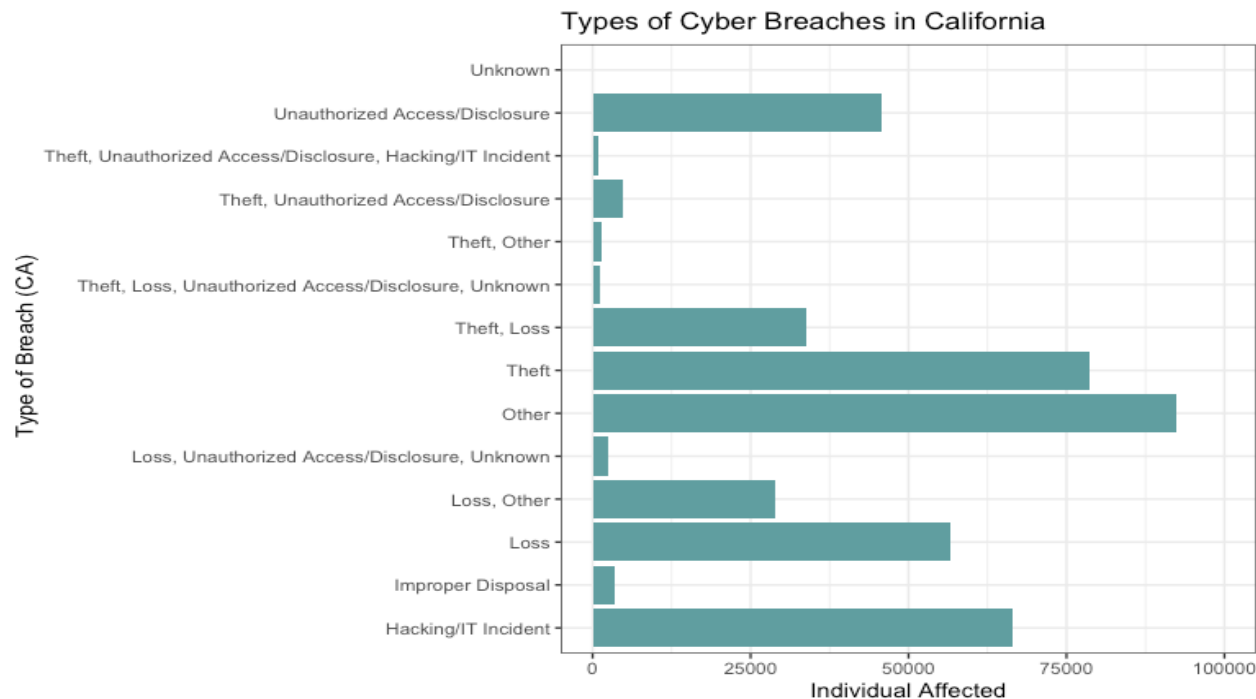> Thus, numbers of people were affected by the breach.

DATA ANALYSIS

Cyber Breaches in USA

As we can see that, apart from all states, the state of California (CA) has accounted highest cyber breaches, where remaining states were also active victim.
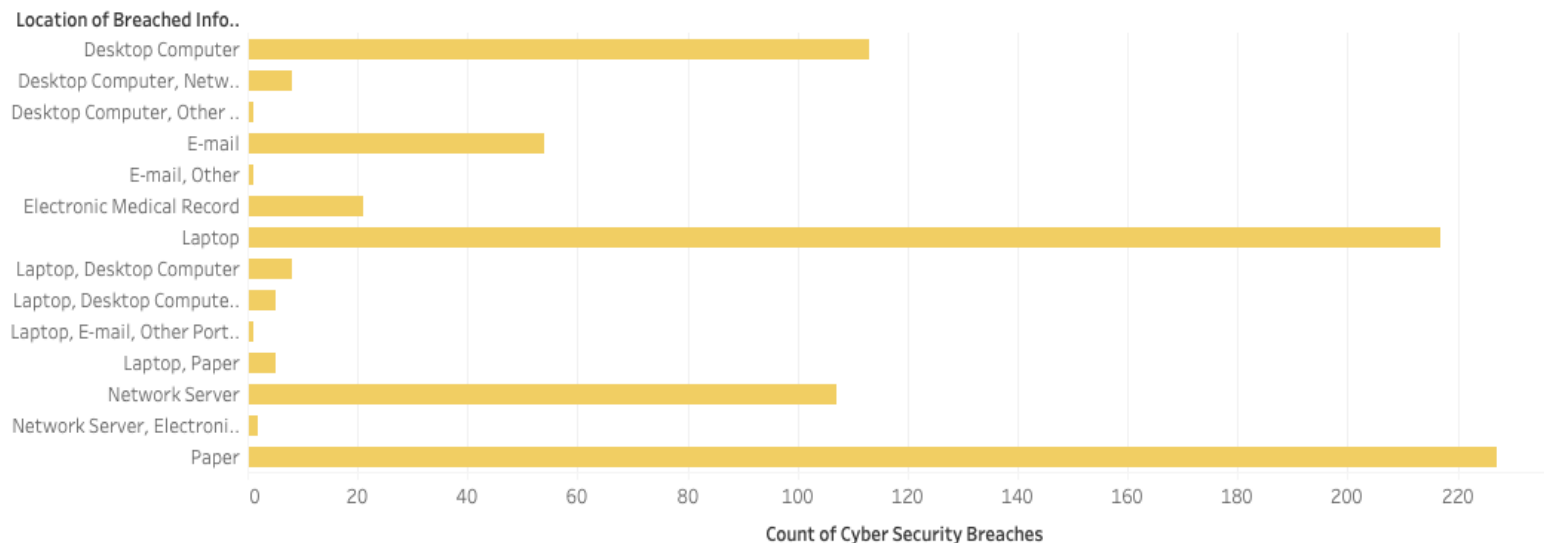
People Affected by Breahes in USA

Hundreds of thousands individuals were affected by the cyber breach in United States Of America. However, less numbers of people suffered from states of Delawer, Hawaii, Maine, and Vermont.
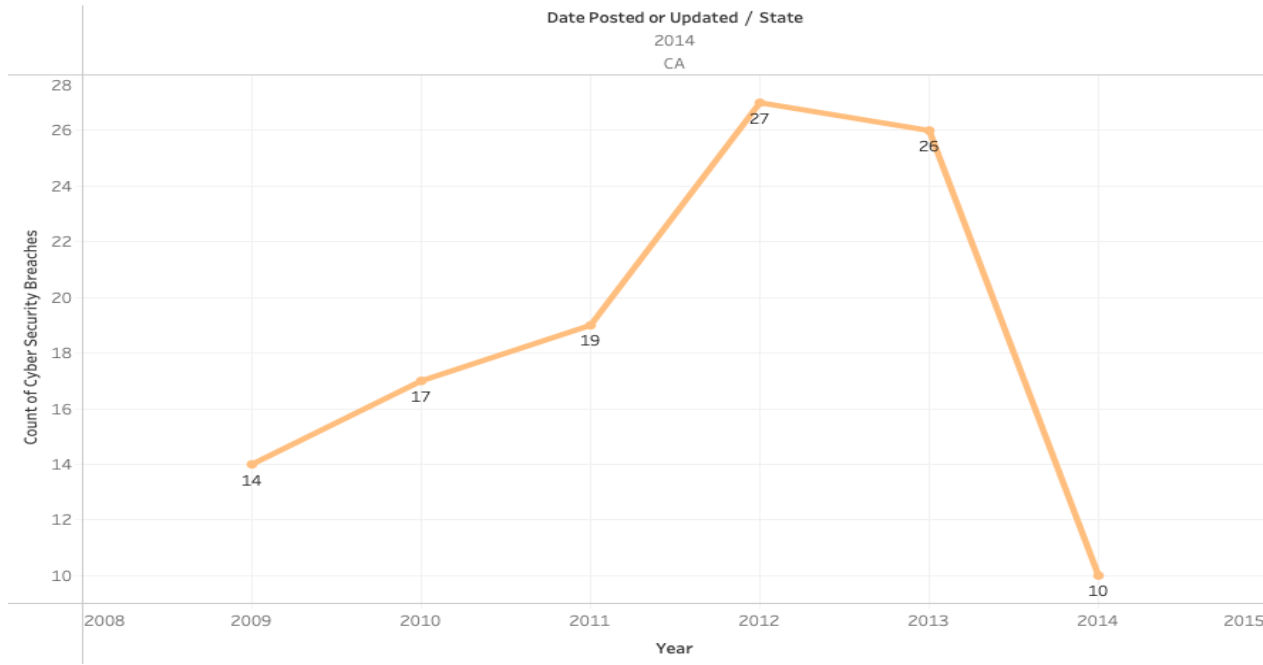
Types of Cyber Breaches in California

Type of Breach (CA):
- Unknown
- Unauthorized Access/Disclosure
- Theft, Unauthorized Access/Disclosure, Hacking/IT Incident
- Theft, Unauthorized Access/Disclosure
- Theft, Other
- Theft, Loss, Unauthorized Access/Disclosure, Unknown
- Theft, Loss
- Theft
- Other
- Loss, Unauthorized Access/Disclosure, Unknown
- Loss, Other
- Loss
- Improper Disposal
- Hacking/IT Incident

Individual Affected (x-axis): 0, 25000, 50000, 75000, 100000

Digging further into the state of California, we can see that bunch of cyber breaches were happened.

Location of Breached Information

Location of Breached Info..:
- Desktop Computer
- Desktop Computer, Netw..
- Desktop Computer, Other ..
- E-mail
- E-mail, Other
- Electronic Medical Record
- Laptop
- Laptop, Desktop Computer
- Laptop, Desktop Compute..
- Laptop, E-mail, Other Port..
- Laptop, Paper
- Network Server
- Network Server, Electroni..
- Paper

Count of Cyber Security Breaches (x-axis): 0, 20, 40, 60, 80, 100, 120, 140, 160, 180, 200, 220
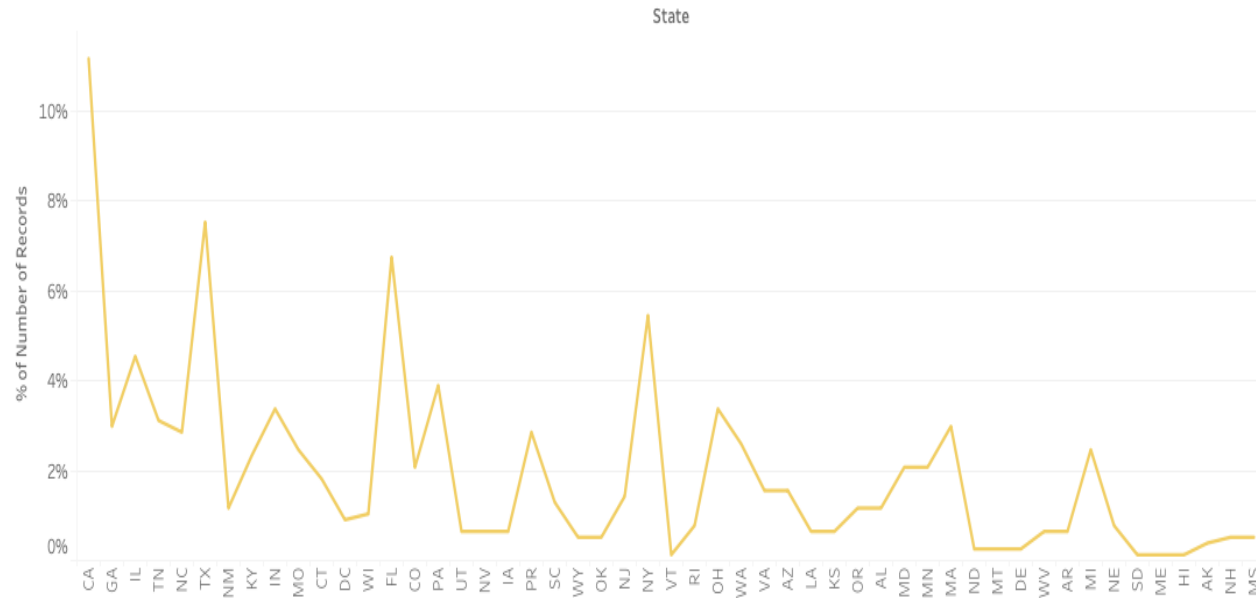
The illustrated chart shows the particular location where the cyber attack happened In California.

Count of Breaches on Yearly Basis in California.

- The line graph, indicates the numbers of cyber breaches happened in California during 2009 to 2014.
- However, in 2012 California reported highest breach *(27)*.

Finally, the line graph illustrates the total number of records had been contributed by each state. *(Out of 1055)*

SUMMARY

By analyzing the dataset, we can say that state of **California** were mostly targeted by hackers.

Numbers of cyber breaches materialized, but **Theft / Unauthorized Access** affected individuals.

Out of all observations, alone California reported **113 records**, and **contribution was 11%.**

In targeted devices **Laptops, Desktops and Confidential papers** had been manipulated.