



CALIFORNIA VOTING TRENDS

ADAM REILLY, CARLEE PRICE, SANJAY DORAIRAJ

GOALS: A LOOK AT CA ELECTION OUTCOME



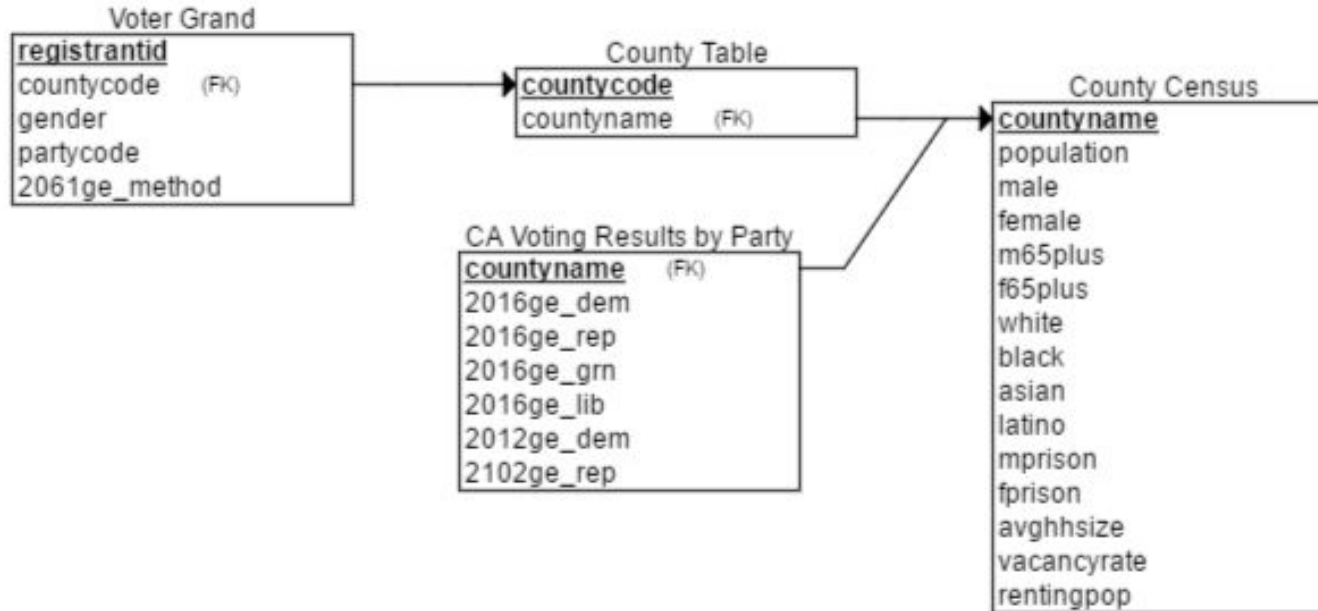
- California is a very blue state -- not everywhere.
- Which counties in the state are outliers (mix of party registrations, turnout)?
- Why do they behave differently (gender mix, socioeconomic issues)?
- In which cases did counties swing (D<>R) or show a narrow margin?
- **What might that mean for future outcomes?**

DATA SOURCES

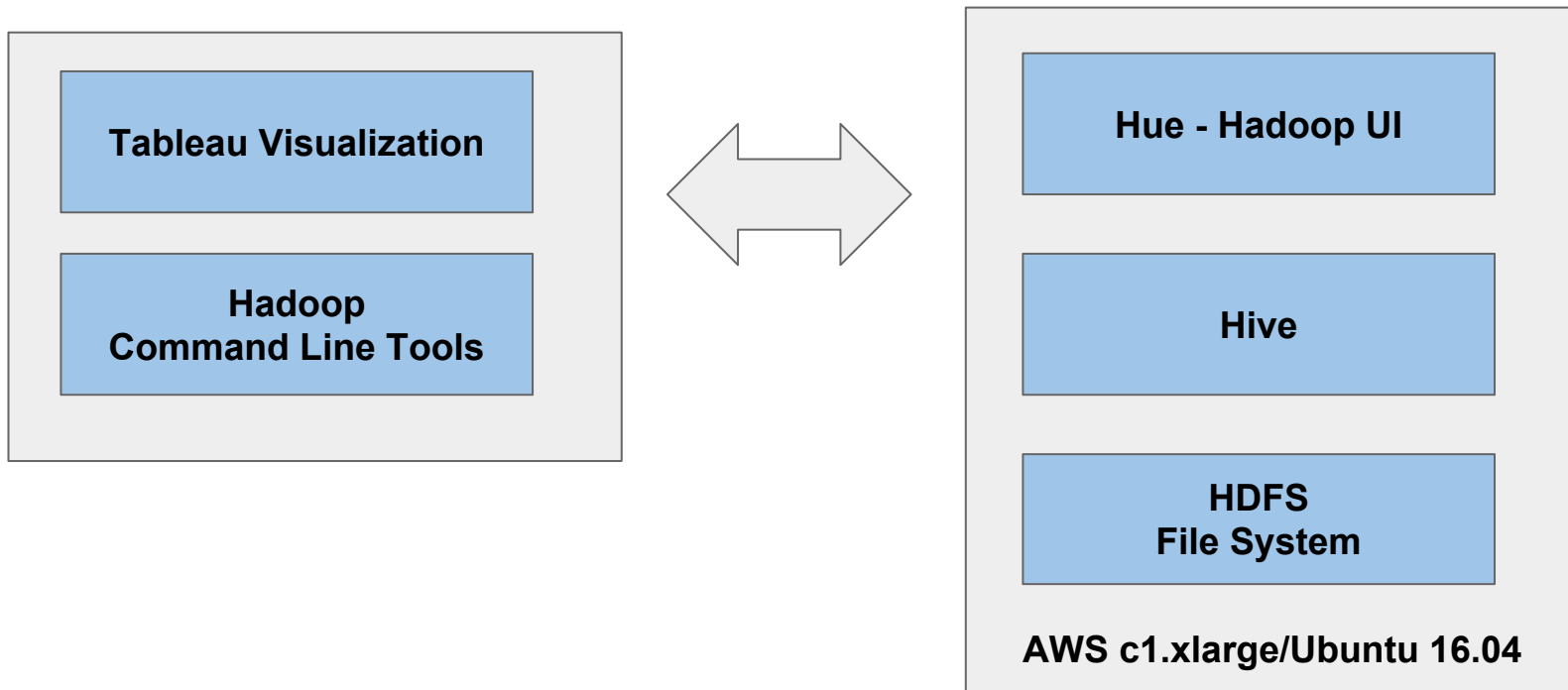


- Election data made available by the State of California Public Voter Registration Data Request (PVRDR)
- County Census data from 2010
- Election outcome data
- Reconcile county codes to county names

RELATIONAL SCHEMA



INFRASTRUCTURE FRAMEWORK



QUERYING PLATFORM



- Hue and Hive were used as the platform for allowing multiple group members to concurrently query the underlying Hive/HDFS data
- Hive metastore backend was changed from SQLite to MySQL in order to address scale issues.

The screenshot shows the Hue web interface. At the top is a navigation bar with links like 'Query Editors', 'Data Browsers', 'Workflows', 'Search', and 'Security'. Below this is a sidebar on the left with a 'Tables' section listing various election-related tables such as '2014ge', '2014ge_1000', '2014ge_10k', '2014pe', '2014pe_1000', '2014pe_10k', '2016ge', '2016ge_1000', '2016pe', '2016pe_1000', '2016pe_10k', 'distinct_voter', 'pvrdr_pd', and 'pvrdr_pd_1000'. The main area is a query editor with a text input field containing an example query: 'SELECT * FROM tablename, or press CTRL + space'. Below the editor is a 'Query History' table with columns for time, status, and query text. The history shows four entries, all with a green checkmark status, representing queries run 'a day ago' and '3 days ago'. The most recent query is a complex one involving multiple tables and a group by clause.

Bookmark Manager Comcast Healthcare Watchable Analytics CCSG - Network O... Google Tools Amazon Web Services UC Berkeley Tech Blogs Health Care Diagno... Machine Learning/AI Health Care Diagno... Comcast

HUE Home Query Editors Data Browsers Workflows Search Security admin

Hive Add a name... Add a description...

< elections (29) ↻ ↺

Tables

- 2014ge
- 2014ge_1000
- 2014ge_10k
- 2014pe
- 2014pe_1000
- 2014pe_10k
- 2016ge
- 2016ge_1000
- 2016pe
- 2016pe_1000
- 2016pe_10k
- distinct_voter
- pvrdr_pd
- pvrdr_pd_1000

1. Example: SELECT * FROM tablename, or press CTRL + space

election:

Query History	Q	Q	Saved Queries
a day ago	✓		select * from pvrdr_vph where electiontype = 'OT' limit 20
a day ago	✓		select * from pvrdr_vph where electiontypedesc = 'OT' limit 20
a day ago	✓		select * from pvrdr_vph limit 20
3 days ago	✓		select registrantid, `2014ge_method`, `2014pe_method`, `2016ge_method`, `2016pe_method` from voter_grand where voter_grand.registrantID in (select registrantID from voter_grand group by registrantID having count(*)

SCOPE: EDA & TRANSFORMATIONS



- *Volume:* Extracted 20GB plus voter registration data into individual components
- *Variety:* Significant cleaning was required (notably gender)
- *Velocity:* Periodic Election Outcomes are easily loaded, as is new census data. By nature, batch rather than streaming.

CHALLENGES



- Data Ingest: file format compatibility, dataset size
- Data Cleaning: duplicates, missing data
- Processing and Serving: OpenRefine, Hadoop
- Structuring queries: Complex SQL to handle state/county comparisons

CALIFORNIA ELECTORAL LANDSCAPE

A TABLEAU STORY BOARD

California Electoral Landscape

<

Party Proportion - Republican

In which counties is the mix of party registrations most

In which counties was turnout (number who voted as a % of

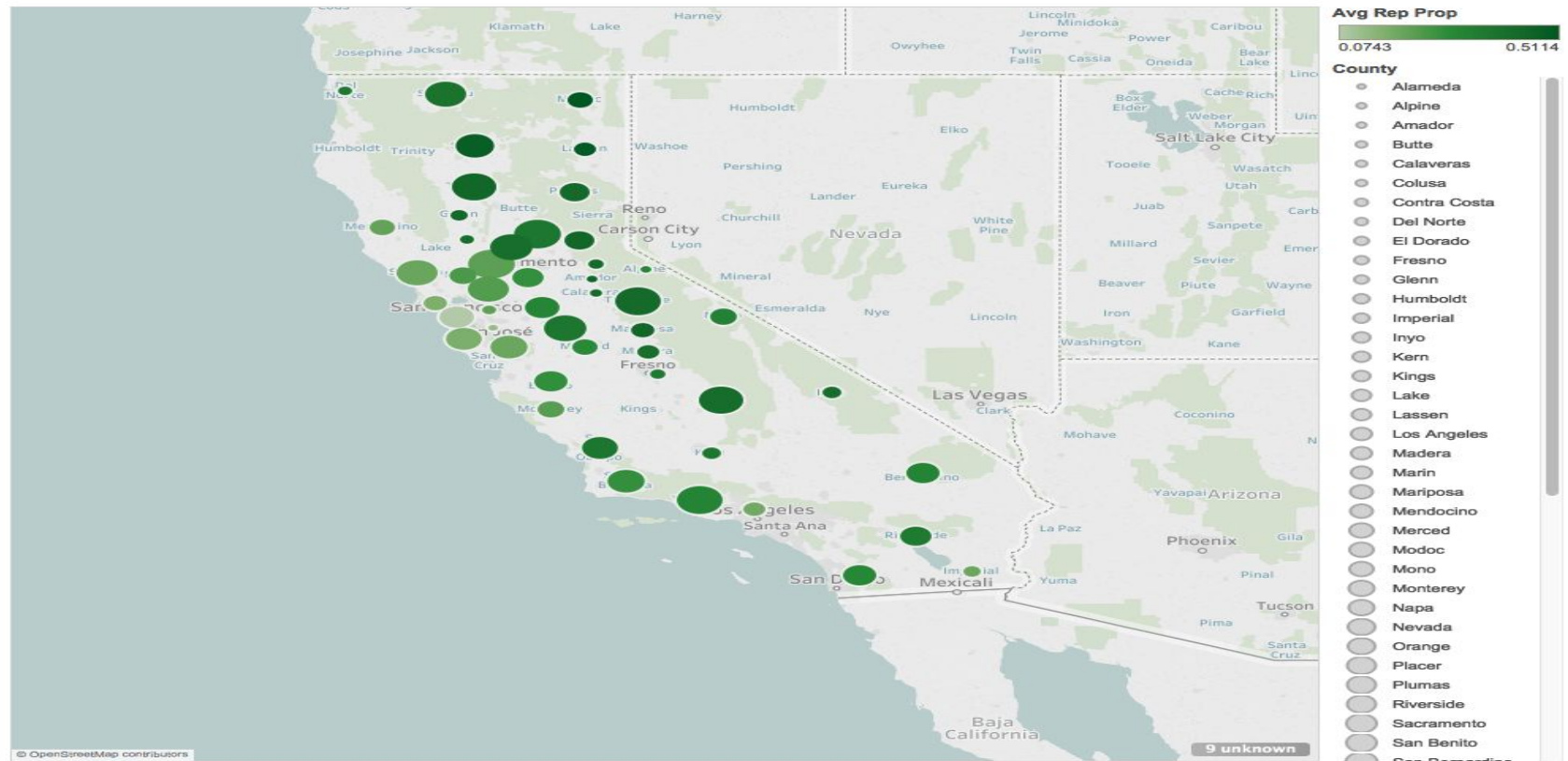
How did statewide turnout (numbers of votes cast compard

What is the breakdown by party affiliation of "first-time

What is the gender breakdown of registrants by party?

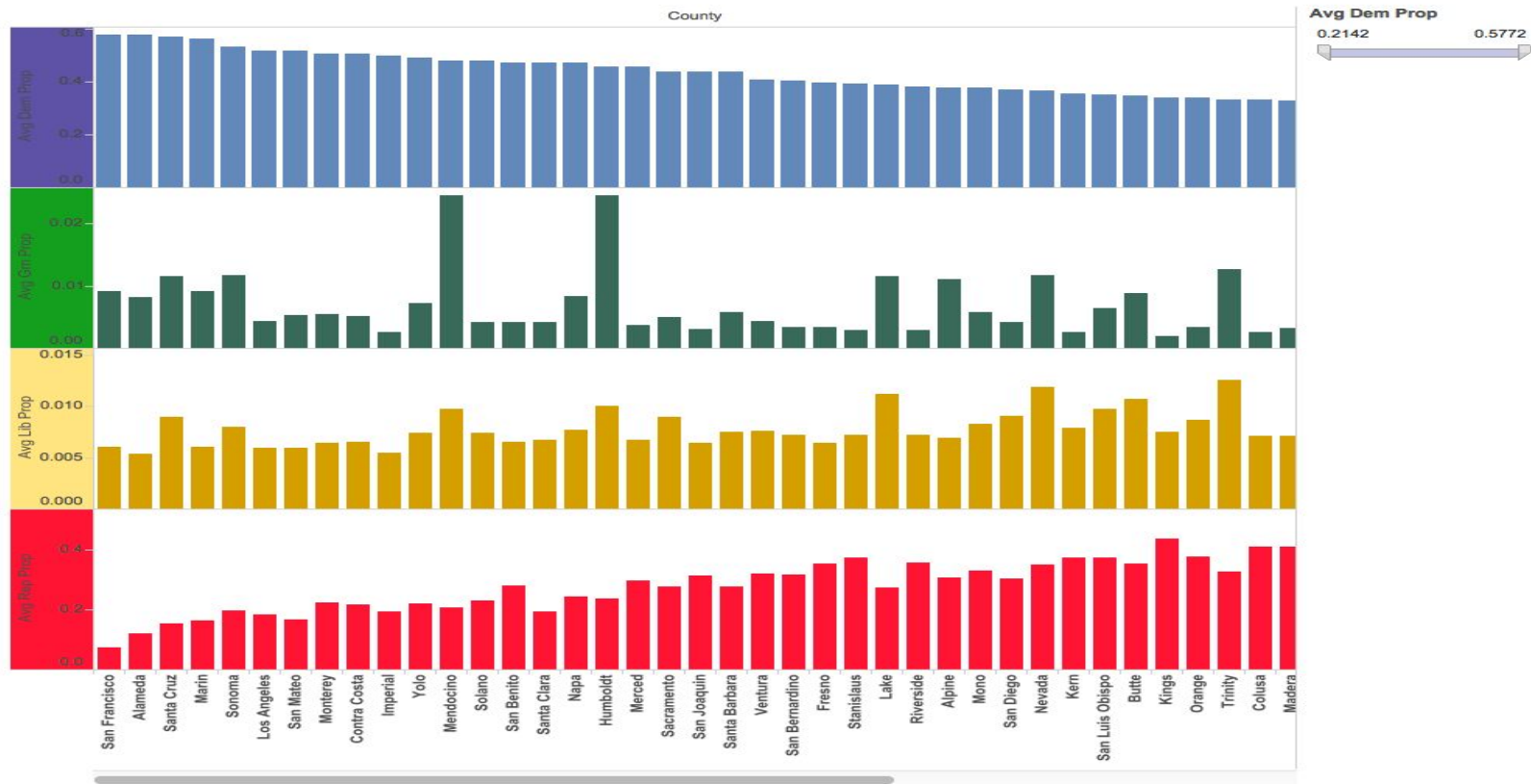
WI ba reg

>



California Electoral Landscape

- <
- Party Proportion - Republican
- In which counties is the mix of party registrations most
- In which counties was turnout (number who voted as a % of
- How did statewide turnout (numbers of votes cast compar
- What is the breakdown by party affiliation of "first-time
- What is the gender breakdown of registrants by party?
- WI ba reg
- >



California Electoral Landscape

<

Party Proportion - Republican

In which counties is the mix of party registrations most

In which counties was turnout (number who voted as a % of

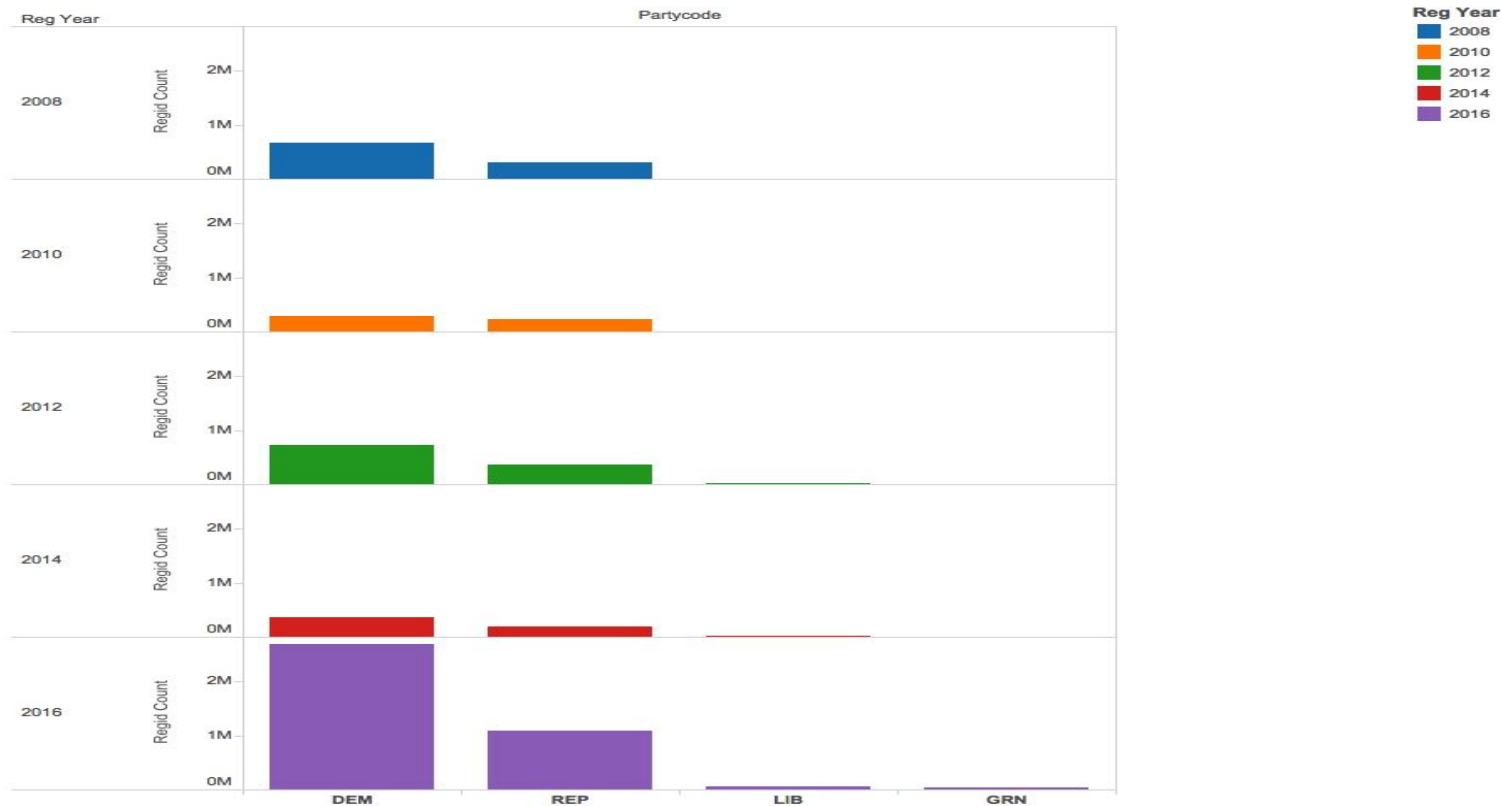
How did statewide turnout (numbers of votes cast compar

What is the breakdown by party affiliation of "first-time

What is the gender breakdown of registrants by party?

WI
ba
reg

>



TRADE-OFFS



- Which research questions yield the best insights
 - Turnout: Registrants v Voters by county
 - Party Mix: by county, gender, first-time voters
 - Party Mix/Turnout: where parties “lost” votes
 - Contested counties
- Which demographic characteristics best highlight outlier counties?
 - Novel insights: Lassen County & prison population
 - Home ownership leads to rightward shift?
 - Likewise high vacancy. Home price appreciation theory for party support?

CONCLUSIONS & LOOKING FORWARD



- Project could easily be scale to look at county-level detail in other states
- Outcome of subsequent elections can be used to test & refine theses



THANK YOU!

PLEASE SEE MORE SLIDES BELOW FOR ADDITIONAL TABLEAU VISUALIZATIONS

California Electoral Landscape

<

Party Proportion - Republican

In which counties is the mix of party registrations most

In which counties was turnout (number who voted as a % of

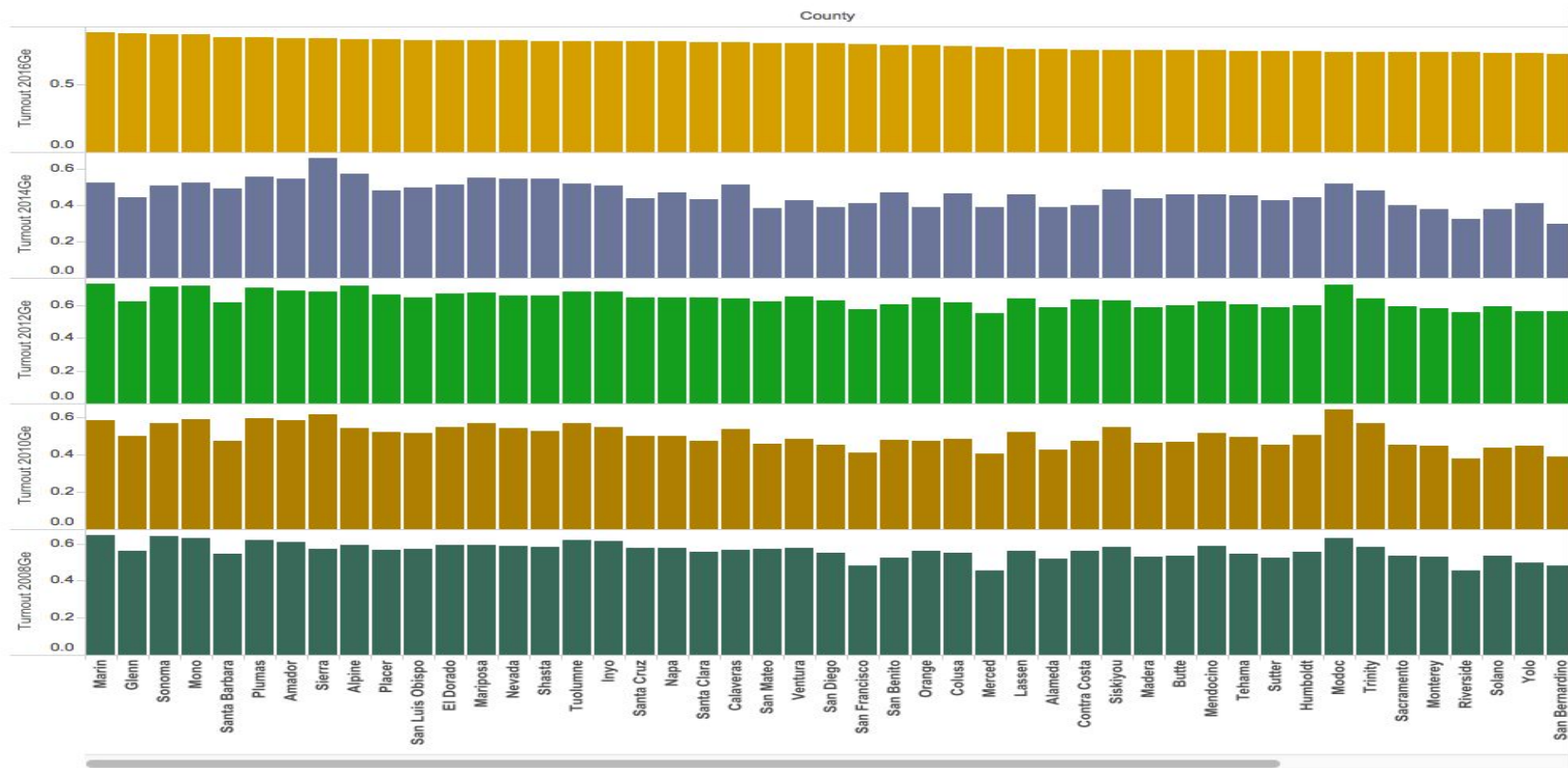
How did statewide turnout (numbers of votes cast compar

What is the breakdown by party affiliation of "first-time

What is the gender breakdown of registrants by party?

What is the gender breakdown of registrants by party?

>



California Electoral Landscape

<

Party Proportion - Republican

In which counties is the mix of party registrations most

In which counties was turnout (number who voted as a % of

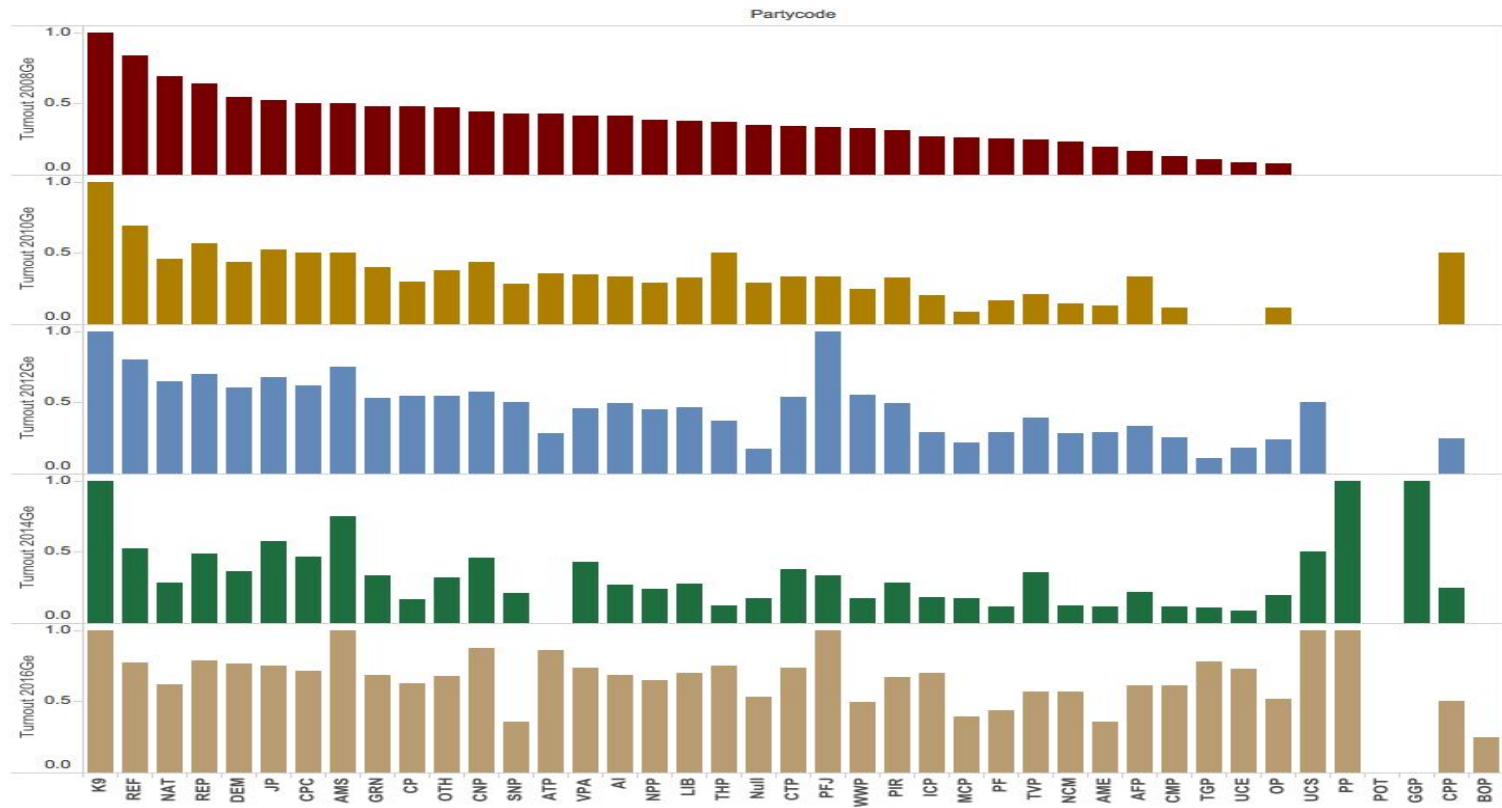
How did statewide turnout (numbers of votes cast compard

What is the breakdown by party affiliation of "first-time

What is the gender breakdown of registrants by party?

WI
ba
reg

>



California Electoral Landscape



Party Proportion -
Republican

In which counties is
the mix of party
registrations most

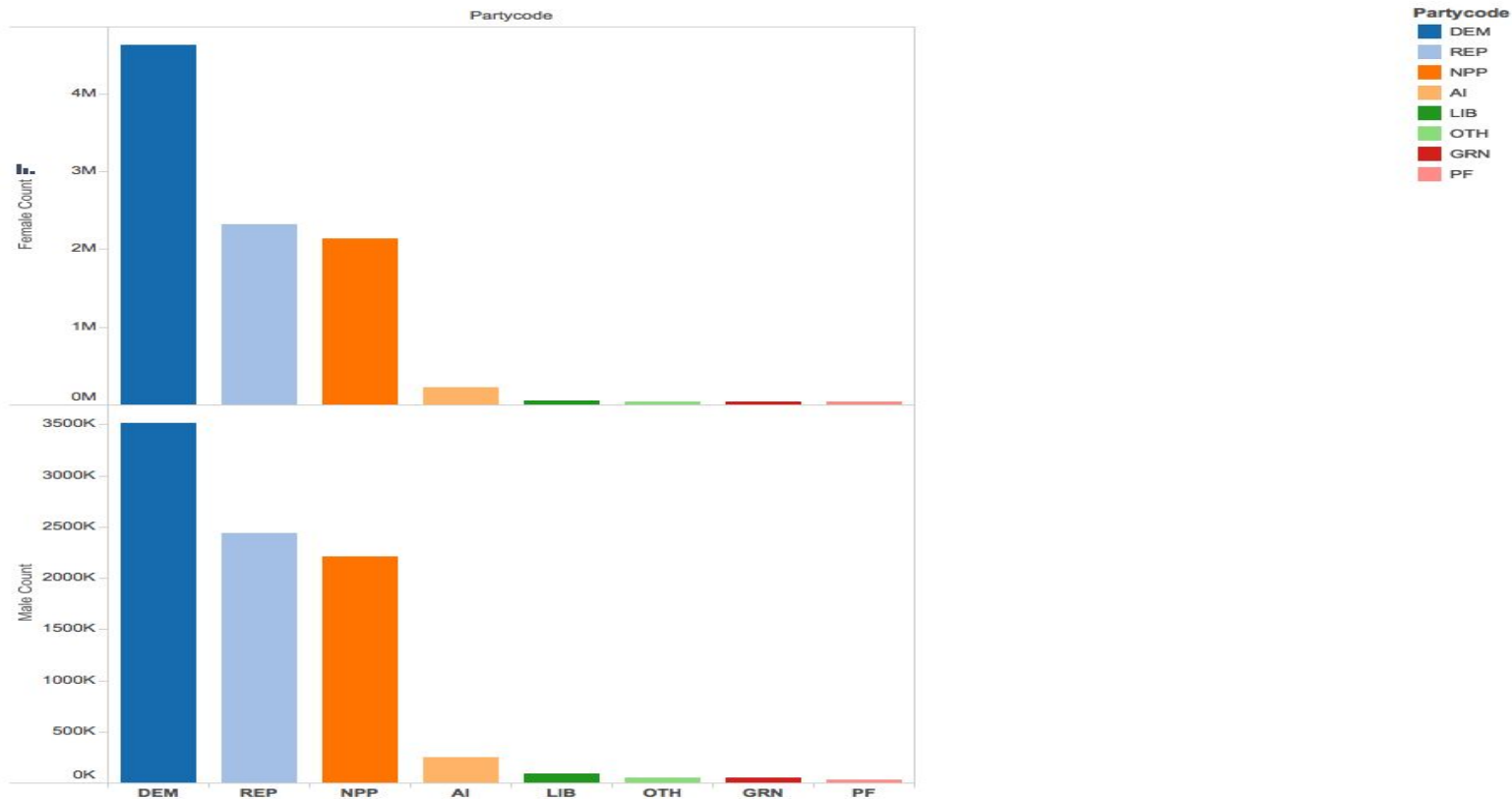
In which counties
was turnout (number
who voted as a % of

How did statewide
turnout (numbers of
votes cast compar

What is the
breakdown by party
affiliation of "first-time

What is the gender
breakdown of
registrants by party?

What is the
breakdown of
registrants by party?



California Electoral Landscape



In which counties is the mix of party registrations most

In which counties was turnout (number who voted as a % of

How did statewide turnout (numbers of votes cast compard

What is the breakdown by party affiliation of "first-time

What is the gender breakdown of registrants by party?

What are gender-based year over year registration trends?

