

**D A N** 

**ВЛИЯНИЕ НА ВЫБОРЫ  
ПРЕЗИДЕНТА США**



# Выборы в США

В США с 1787 года на выборах президента действует двухпартийная система

- Представлены Демократическая и Республиканская партии
- Выборы проходят каждые 4 года
- Выборы проходят по штатам
- Голосование проходит в 2 этапа

X



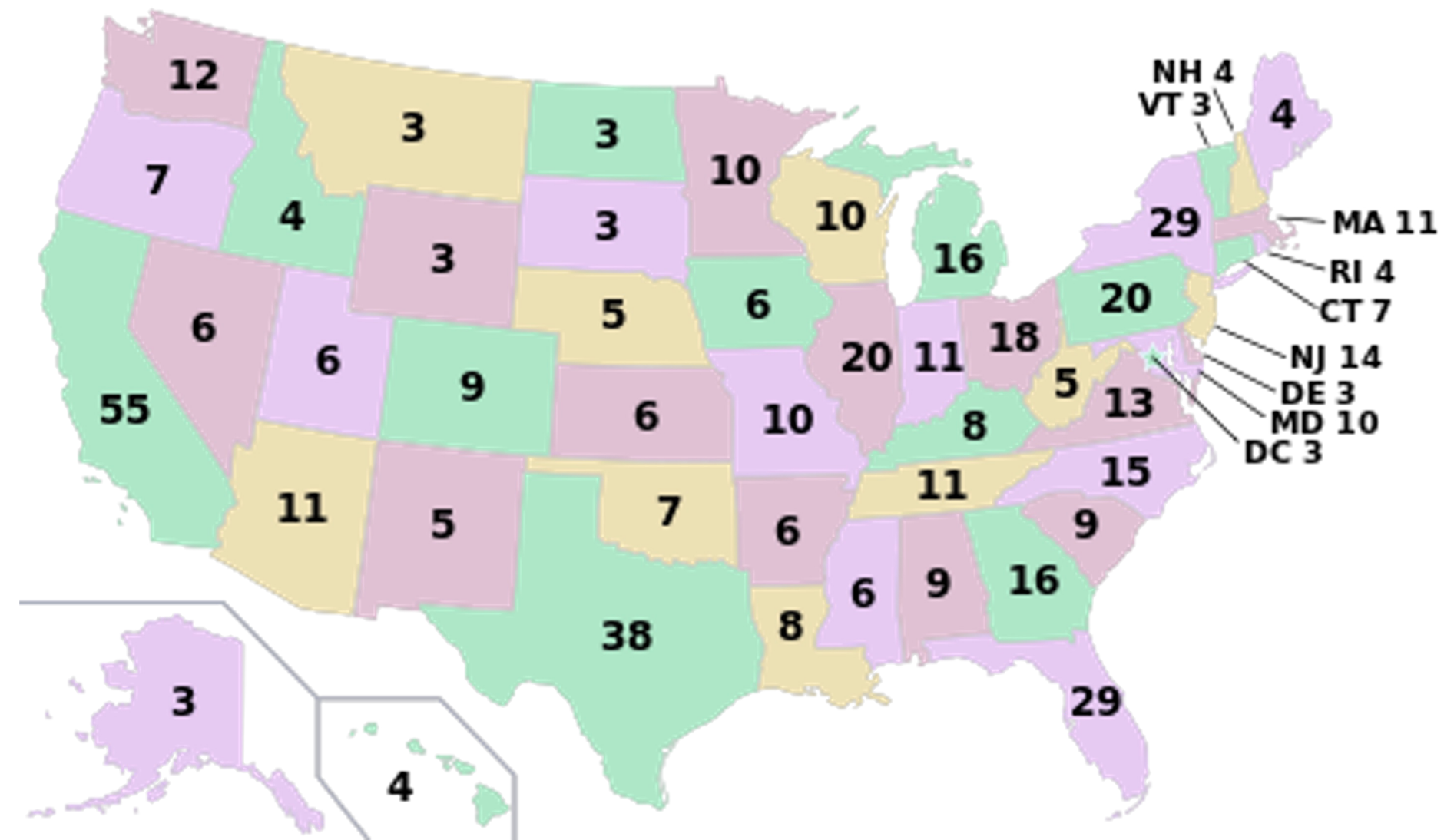
# Этап 1 - выбор «выборщиков»



Совершеннолетние граждане штата выбирают один из двух списков выборщиков. Один список - республиканской партии, второй - Демократической

В каждом штате разное количество выборщиков. Чем больше население штата, тем больше выборщиков.

Всего выборщиков **538**





# Описание данных



Вам предоставлена статистика по населению округов каждого штата за 2009-2013 годы

**В данных есть характеристики населения по:**

- возрасту
- расе
- образованию
- стоимости жилья
- доходу
- количеству бизнесов разных отраслей
- различию количества голосов за партии



Описание переменных

# Округ

- **area\_name** - Название округа
- **state\_abbrev** - Аббревиатура – обозначение штата



Описание переменных

# Демография

- **PST045214** - Численность населения округа на 2014 год
- **AGE135214** - Процент детей младше 5ти лет, 2014 год
- **AGE295214** - Процент детей до 18ти лет, 2014 год
- **AGE775214** - Процент людей старше 65лет, 2014 год
- **SEX255214** - Кол-во женщин, 2014 год
- **RHI125214** - Количество белого населения, 2014 год
- **RHI225214** - Процент афроамериканцев, 2014 год
- **RHI325214** - Процент коренного населения (индейцы и коренные жители Аляски), 2014 год
- **RHI425214** - Процент азиатов, 2014



Описание переменных

# Демография

- **RHI525214** - Процент коренного населения Гавайских островов и других Тихоокеанских островов, 2014
- **RHI625214** - Процент метисов (предки двухи и более рас), 2014
- **RHI725214** - Процент испанского и латиноамериканского населения, 2014
- **RHI825214** - Процент белого населения (не испанцы и не латиноамериканцы), 2014
- **POP645213** - Процент людей, рожденных не в США, 2009-2013
- **VET605213** - Численность ветеранов войны, 2009-2013
- **POP815213** - Процент детей до 5-ти лет, у которых дома говорят более чем на одном языке, 2009-2013
- **EDU635213** - Процент людей старше 25ти окончивших как минимум старшую школу, 2009-2013
- **EDU685213** - Процент людей старше 25ти, получивших минимум степень бакалавра, 2009-2013





Описание переменных

# Жилье и благосос- тояние

- **LFE305213** - Среднее время в минутах на дорогу от дома для работы (данные по работникам старше 16ти лет), 2009-2013
- **HSG010214** - Число жилых помещений, 2014
- **HSG445213** - Уровень домовладения (какой процент людей владеет собственной недвижимостью), 2009-2013
- **HSG096213** - Процент жилых помещений в многоквартирных домах, 2009-2013
- **HSD410213** - Количество домашних хозяйств, 2009-2013
- **HSD310213** - Сколько человек приходится на одно домашнее хозяйство, 2009-2013
- **INC910213** - Денежный доход на душу населения за последние 12 месяцев (в долларах 2013 г.), 2009-2013
- **INC110213** - Медианный доход домашних хозяйств, 2009-2013



Описание переменных

# Жилье и благосос- тояние

- **PVY020213** - Процент людей находящихся ниже уровня бедности, 2009-2013
- **BZA010213** - Количество частных несельскохозяйственных предприятий, 2013
- **BZA110213** - Количество занятых на частных несельскохозяйственных предприятиях, 2013
- **BZA115213** - Процентное изменение количества занятых на частных несельскохозяйственных предприятиях, 2012-2013
- **NES010213** - Кол-во предприятий не являющихся работодателями, 2013
- **BPS030214** - Кол-во выданных разрешений на строительство, 2014
- **LND110210** - Площадь округа в квадратных милях, 2010



Описание переменных

# Результаты выборов

- **Demvotes16** - Число голосов за кандидата от Демократической партии, 2016
- **GOPvotes16** - Число голосов за кандидата от Республиканской партии, 2016
- **total\_2016** - Общее число голосов в 2016м году
- **pct\_dem\_16** - Процент голосов за Демократическую партию от общего числа голосов, 2016
- **pct\_gop\_16** - Процент голосов за Республиканскую партию от общего числа голосов, 2016
- **diff\_2016** - Разница в числе голосов между Республиканской и Демократической партией
- **pct\_pt\_16** - Разница ( $diff\_2016$ ) в процентах от общего числа. Отрицательная если за Демократическую партию отдали меньше голосов.

## Исследовательский вопрос

X

Каким образом социально-демографическая и экономическая ситуация в стране повлияла на результаты выборов?

# Какие факторы повлияли

**Какие признаки можно использовать для анализа причины различия мнений?**

Например, для одного округа будет более значимо расовое рассмотрение, а для другого - неравенство в доходах.

**Какие характеристики избирателей влияют на решение о голосовании?**

Например, где-то влияет раса, а где-то достаток

**По-разному ли голосуют разные округа?**

**Что может быть причиной различий?**

# Какие факторы повлияли

**Какое распределены голоса? Есть ли какие-то особенности в этом распределении?**

Попробуйте построить график и посмотреть на разницу в голосах среди разных регионов

**Как можно классифицировать округа по результатам выборов президента США?**

Есть ли какие-то группы округов с похожими результатами голосования?

**Как можно использовать данные о результатах выборов президента США по округам для прогнозирования будущих выборов?**

На какие социально-демографические группы тогда стоит ориентироваться партиям при выстраивании предвыборной кампании?

X

## Какие зависимости вы ожидаете?

- чем старше население, тем больше голосов за консервативную партию
- чем больше расовое разнообразие округа, тем больше голосов за либеральную партию
- чем больше средний доход, тем больше голосов за консервативную партию
- чем более развит бизнес, тем больше голосов за консервативную партию



X

## Агрегирование данных

Это необходимо, чтобы исследовать связь между характеристиками населения и округами

Как агрегировать? (по округу, по году выборов, ...)







# Как классифицировать округа?

Как предсказать решение жителей до завершения выборов?





## Анализ

Учитывая все вышеперечисленное, проведите анализ данных и выявите ключевые факторы, влияющие на выбор кандидата в президенты в округе





# **Выводы и практические рекомендации**

## **Подведите итог проведенному исследованию**

- Что именно вы хотели выяснить?
- Какие результаты были получены?
- Как эти результаты могут быть использованы на практике?
- Какие ограничения своего исследования вы видите?
- Если бы у вас было больше времени и не было ограничений, связанных с доступностью данных, как вы могли бы дополнить и улучшить свое исследование?

# Помните

X



Представленные ранее пункты являются примерными и отражают **«программу максимум»**.



Успешная работа не обязательно требует выполнения всех из них — вы можете сфокусироваться **на отдельных аспектах**.



Более того — работа, глубоко исследующая какой-то один важный аспект представляет **большую ценность**, чем работа, исследующая все, но поверхностно.

**DAN** 

**СПАСИБО**

