

Для цього проекту я обрала цю базу даних, <https://www.coshnetwork.org/fatality-database>, бо вона містить різноманітну інформацію про нещасні випадки на робочих місцях. Я хочу дослідити у ній залежність між смертями та такими характеристиками як вік та стать по роках та штатах за останні роки, а також, як Covid-19 вплинув на цю статистику.

Перше, що я зробила - це знайшла статистику за гендером по штатах за один рік, але потім зрозуміла, що інформація сильно варіюється і не дає жодної загальної картини (рис. 0), тому я прийняла рішення відразу дослідити цю інформацію за всі роки (рис. 1). Найкраще для такого типу візуалізацій підходить stacked bar chart, який показує абсолютне значення та дає змогу порівняти інформацію за гендером/віком.

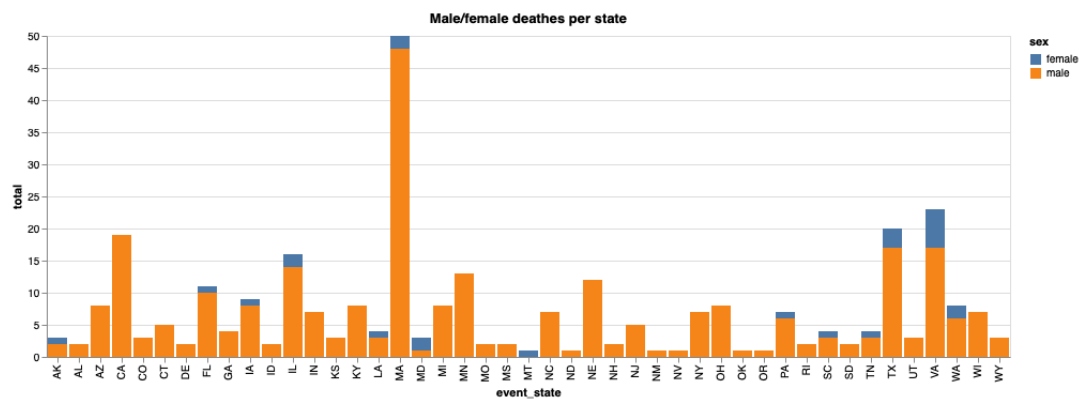


Рисунок 0

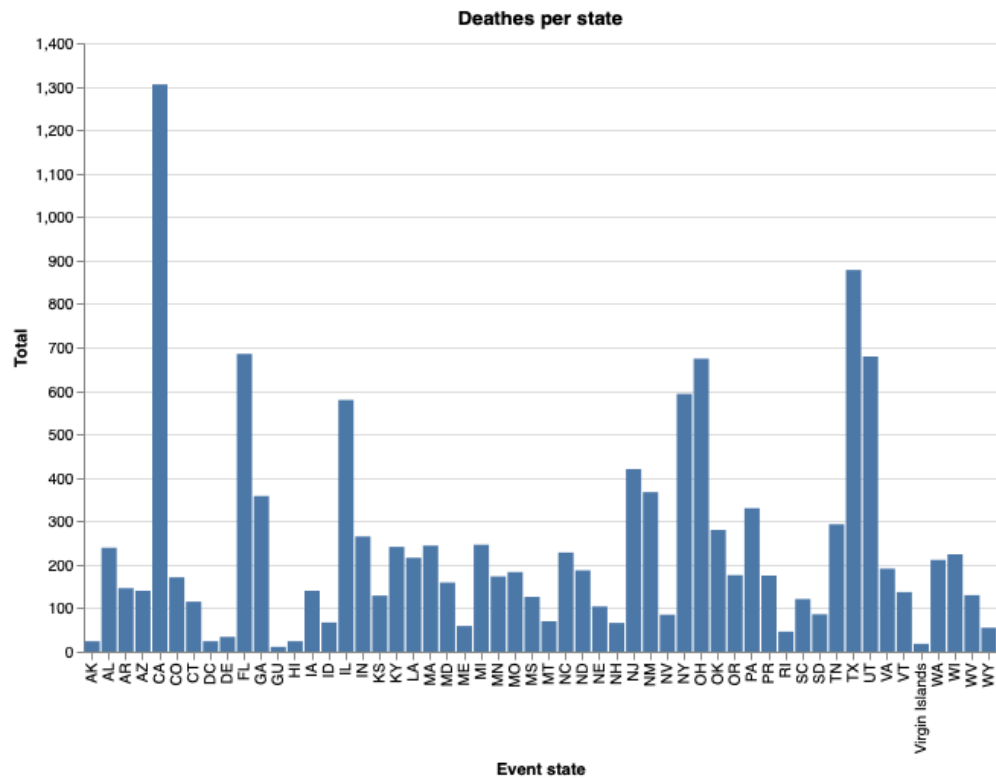


Рисунок 1

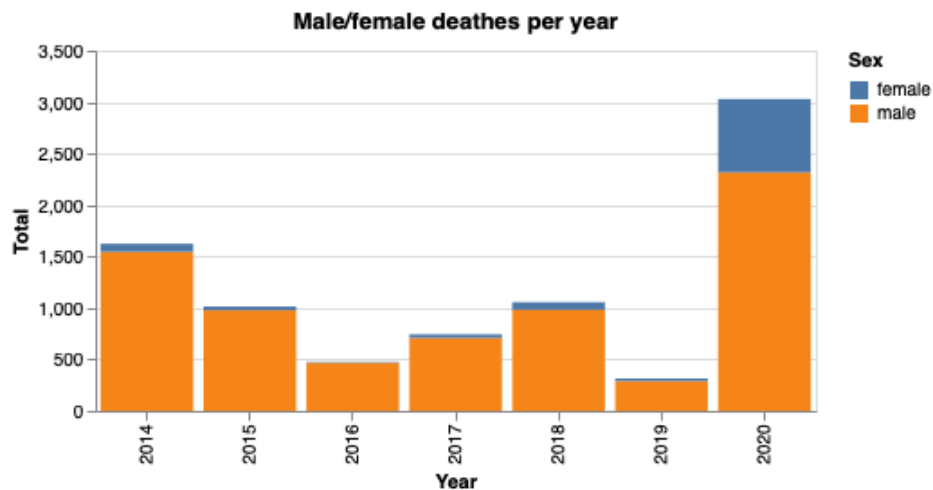


Рисунок 2

Висновок: інформація досі сильно варіюється, проте тепер це підтверджено роками, що у деяких штатах люди гинуть частіше на робочих місцях. Також очевидно із обох графіків, що чоловіки стають жертвами нещасних випадків в рази частіше, ніж жінки, хоч в деяких штатах помітно відмінності.

Потім мені стало цікаво дослідити цю статистику щодо віку і так як вік сильно варіювався, я вирішила поділити його на проміжки по 5 років, щоб краще зрозуміти, яка вікова категорія знаходиться під найбільшим ризиком.

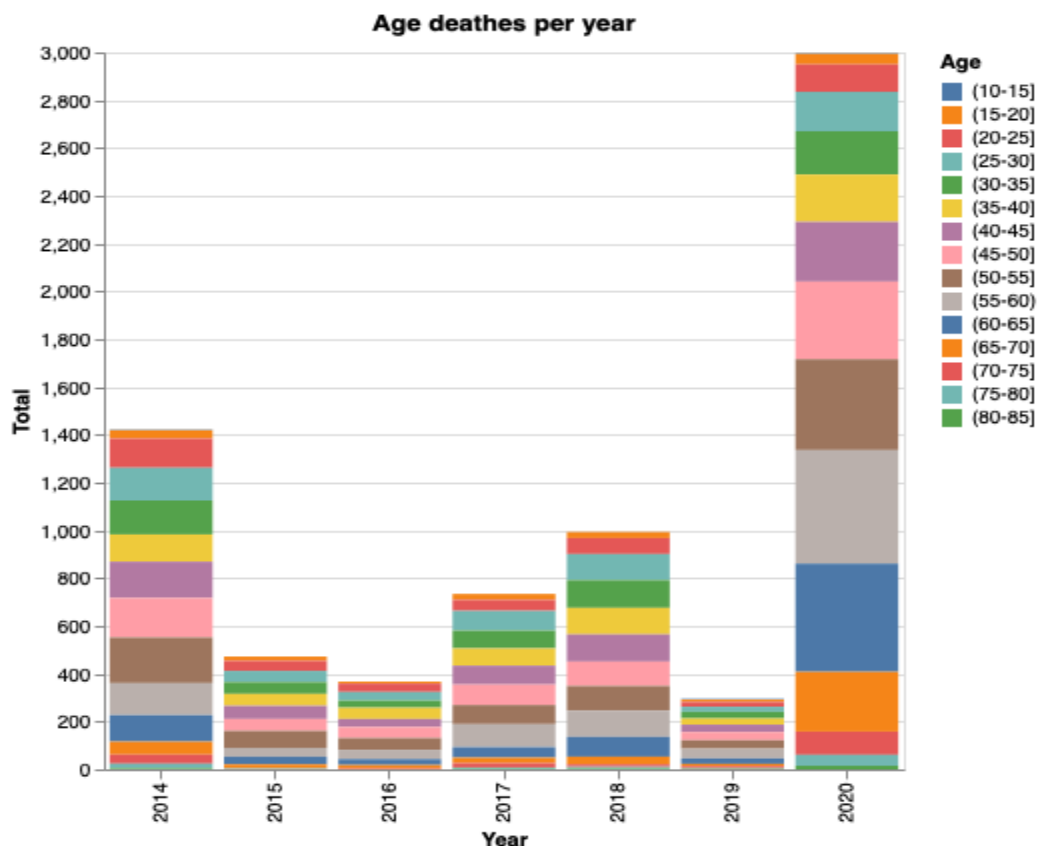


Рисунок 3

Висновок: передбачувано, пенсійна та передпенсійна категорія в найбільшому ризику, але представники усіх вікових категорій можуть стати жертвою нещасного випадку на робочому місці.

Що найбільше кидається в очі на другому та третьому рисунках - великий скачок смертей, особливо представників жіночої статі, у 2020 році. Спочатку я подумала, що це пов'язано з вірусом COVID-19, проте у файлі тільки 10% смертей були пов'язані з ним, тому я вирішила зробити графік смертей протягом року. Для цієї візуалізації я обрала bar chart, щоб найкраще показати "екстремуми" інформації, не переймаючись, що лінія буде кривою і "стрибати" між суттєво різними значеннями.

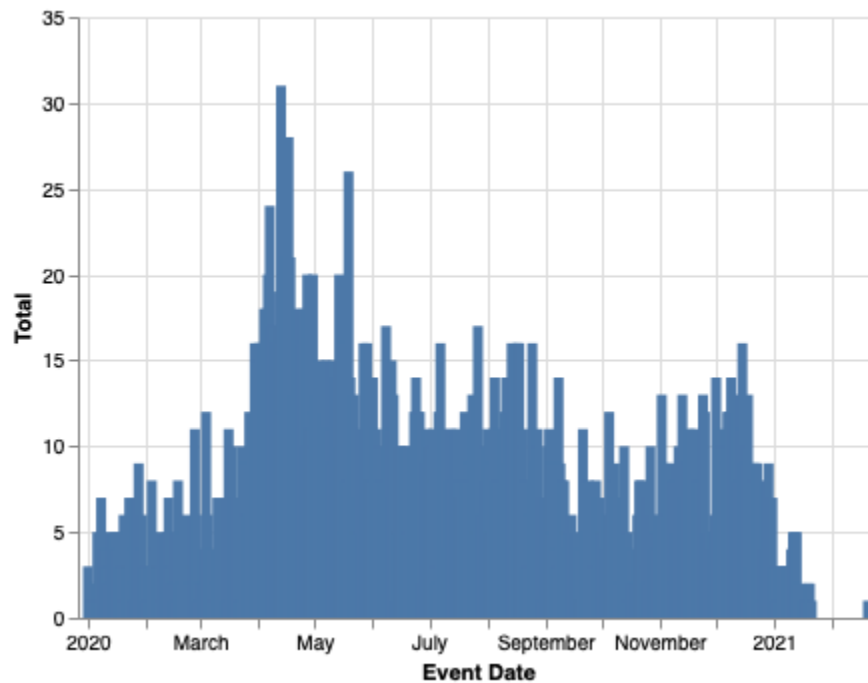


Рисунок 4

Висновок: Це було правильним рішенням, бо пік графіку(рисунок 4) показав перший пік коронавірусу, що свідчить про неточність колонки “причина смерті” у ексель файлі.

Загальні висновки:

- Чоловіки набагато частіше помирають на робочому місці
- Каліфорнія, Техас та Флорида є трійкою рекордсменів по нещасних випадках за 2014-2020pp
- Хоч і люди пенсійного віку більш схильні стати жертвою на роботі, загальна статистика вказує, що великої різниці між віковими категоріями у даному питанні немає
- Скачок смертей у 2020 (і відсотковий скачок у смертях жінок) році зумовлений пандемією коронавірусної хвороби