

Авторы датасета

The Centers for Disease Control and Prevention (CDC)
doi: 10.24432/C53919

Ссылка на оригинал датасета: <https://www.kaggle.com/datasets/cdc/behavioral-risk-factor-surveillance-system>

Версия датасета

14 марта 2023

Описание датасета

Файл «diabetes_BRFSS2015.csv» состоит из данных 253.680 пациентов с подозрением на диабет, опрошенных в рамках исследования CDC BRFSS в 2015 году. Датасет содержит 22 атрибута для анализа рисков развития сахарного диабета.

BRFSS (Система надзора за поведенческими факторами риска) представляет собой телефонный опрос, связанный со здоровьем, который ежегодно проводится в США организацией CDC.

Детальное описание полей в датасете:

- **Diabetes_binary** — Целевая переменная наличия сахарного диабета (0 = нет диабета; 1 = преддиабет; 2 = диабет)
- **HighBP** — Артериальное давление (0 = низкое; 1 = высокое)
- **HighChol** — Уровень холестерина (0 = низкий; 1 = высокий)
- **CholCheck** — Проверялся ли уровень холестерина в течение последних 5 лет? (0 = нет; 1 = да)
- **BMI** — Индекс массы тела
- **Smoker** — Выкурили ли вы за всю свою жизнь хотя бы 100 сигарет? (0 = нет; 1 = да)
- **StrokeFeature** — Был ли у вас когда-нибудь зафиксирован инсульт? (0 = нет; 1 = да)
- **HeartDiseaseorAttack** — Были ли у вас когда-нибудь зафиксированы болезни сердца или инфаркт? (0 = нет; 1 = да)
- **PhysActivity** — Занимались ли вы в последние 30 дней активными физическими упражнениями, не считая работы? (0 = нет; 1 = да)
- **Fruits Feature** — Употребляете ли вы регулярно фрукты один или более раз в день? (0 = нет; 1 = да)
- **Veggies** — Употребляете ли вы регулярно овощи один или более раз в день? (0 = нет; 1 = да)
- **HvyAlcoholConsump** — Употребляете ли вы крепкие спиртные напитки: мужчины 14 порций в неделю или чаще; женщины 7 порций в неделю или чаще? (0 = нет; 1 = да)
- **AnyHealthcare** — Есть ли у вас какая-либо медицинская страховка? (0 = нет; 1 = да)

- **NoDocbcCost** — Были ли за последние 12 месяцев случаи, когда вы бы хотели попасть на прием к доктору, но не могли финансово себе это позволить? (0 = нет; 1 = да)
- **GenHlth** — Как бы вы в общем оценили уровень вашего здоровья? (1 = прекрасное; 2 = очень хорошее; 3 = хорошее; 4 = удовлетворительное; 5 = плохое)
- **MentHlth** — Если говорить о вашем психическом здоровье, имея в виду депрессию, стресс, эмоциональное выгорание и пр., то за последние 30 дней сколько дней ваше психическое здоровье было неудовлетворительным? (ответ: количество дней от 0 до 30)
- **PhysHlth** — Если говорить о вашем физическом здоровье, имея в виду болезни, травмы и пр., то за последние 30 дней сколько дней ваше физическое здоровье было неудовлетворительным? (ответ: количество дней от 0 до 30)
- **DiffWalk** — Возникают ли у вас серьезные трудности при длительных пешеходных прогулках или подъеме по лестнице? (0 = нет; 1 = да)
- **Sex** — Пол (0 = женский; 1 = мужской)
- **Age** — Возрастная категория (1 = 18-24; 2 = 25-29; 3 = 30-34; 4 = 35-39; 5 = 40-44; 6 = 45-49; 7 = 50-54; 8 = 55-59; 9 = 60-64; 10 = 65-69; 11 = 70-74; 12 = 75-79; 13 = 80 и старше)
- **Education** — Уровень образования (1 = Никогда не посещал школу, только детский сад; 2 = Начальное общее образование; 3 = Основное общее образование; 4 = Среднее общее образование; 5 = Среднее профессиональное образование или Бакалавр; 6 = Магистр)
- **Income** — Доход (1 = менее 10 тыс. долларов в год; 2 = менее 15 тыс. долларов в год; 3 = менее 20 тыс. долларов в год; 4 = менее 25 тыс. долларов в год; 5 = менее 35 тыс. долларов в год; 6 = менее 50 тыс. долларов в год; 7 = менее 75 тыс. долларов в год; 8 = более 75 тыс. долларов в год)