**About Dataset**

This dataset contains comprehensive health data for 1,879 patients, uniquely identified with IDs ranging from 6000 to 7878. The data includes demographic details, lifestyle factors, medical history, clinical measurements, medication usage, symptoms, quality of life scores, environmental exposures, and health behaviors. Each patient is associated with a confidential doctor in charge, ensuring privacy and confidentiality. This dataset is ideal for researchers and data scientists looking to explore factors associated with diabetes, develop predictive models, and conduct statistical analyses.

**Table of Contents**

1. **Patient Information**

* Patient ID
* Demographic Details
* Lifestyle Factors

1. **Medical History**
2. **Clinical Measurements**
3. **Medications**
4. **Symptoms and Quality of Life**
5. **Environmental and Occupational Exposures**
6. **Health Behaviors**
7. **Diagnosis Information**

**Patient Information**

**Patient ID**

* **PatientID**: A unique identifier assigned to each patient (6000 to 7878).

**Demographic Details**

* **Age**: The age of the patients ranges from 20 to 90 years.
* **Gender**: Gender of the patients, where 0 represents Male and 1 represents Female.
* **Ethnicity**: The ethnicity of the patients, coded as follows:
  + 0: Caucasian
* 1: African American
* 2: Asian
* 3: Other
* **SocioeconomicStatus**: The socioeconomic status of the patients, coded as follows:
  + 0: Low
* 1: Middle
* 2: High
* **EducationLevel**: The education level of the patients, coded as follows:
  + 0: None
* 1: High School
* 2: Bachelor's
* 3: Higher

**Lifestyle Factors**

* **BMI**: Body Mass Index of the patients, ranging from 15 to 40.
* **Smoking**: Smoking status, where 0 indicates No and 1 indicates Yes.
* **AlcoholConsumption**: Weekly alcohol consumption in units, ranging from 0 to 20.
* **PhysicalActivity**: Weekly physical activity in hours, ranging from 0 to 10.
* **DietQuality**: Diet quality score, ranging from 0 to 10.
* **SleepQuality**: Sleep quality score, ranging from 4 to 10.

**Medical History**

* **FamilyHistoryDiabetes**: Family history of diabetes, where 0 indicates No and 1 indicates Yes.
* **GestationalDiabetes**: History of gestational diabetes, where 0 indicates No and 1 indicates Yes.
* **PolycysticOvarySyndrome**: Presence of polycystic ovary syndrome, where 0 indicates No and 1 indicates Yes.
* **PreviousPreDiabetes**: History of previous pre-diabetes, where 0 indicates No and 1 indicates Yes.
* **Hypertension**: Presence of hypertension, where 0 indicates No and 1 indicates Yes.

**Clinical Measurements**

* **SystolicBP**: Systolic blood pressure, ranging from 90 to 180 mmHg.
* **DiastolicBP**: Diastolic blood pressure, ranging from 60 to 120 mmHg.
* **FastingBloodSugar**: Fasting blood sugar levels, ranging from 70 to 200 mg/dL.
* **HbA1c**: Hemoglobin A1c levels, ranging from 4.0% to 10.0%.
* **SerumCreatinine**: Serum creatinine levels, ranging from 0.5 to 5.0 mg/dL.
* **BUNLevels**: Blood Urea Nitrogen levels, ranging from 5 to 50 mg/dL.
* **CholesterolTotal**: Total cholesterol levels, ranging from 150 to 300 mg/dL.
* **CholesterolLDL**: Low-density lipoprotein cholesterol levels, ranging from 50 to 200 mg/dL.
* **CholesterolHDL**: High-density lipoprotein cholesterol levels, ranging from 20 to 100 mg/dL.
* **CholesterolTriglycerides**: Triglycerides levels, ranging from 50 to 400 mg/dL.

**Medications**

* **AntihypertensiveMedications**: Use of antihypertensive medications, where 0 indicates No and 1 indicates Yes.
* **Statins**: Use of statins, where 0 indicates No and 1 indicates Yes.
* **AntidiabeticMedications**: Use of antidiabetic medications, where 0 indicates No and 1 indicates Yes.

**Symptoms and Quality of Life**

* **FrequentUrination**: Presence of frequent urination, where 0 indicates No and 1 indicates Yes.
* **ExcessiveThirst**: Presence of excessive thirst, where 0 indicates No and 1 indicates Yes.
* **UnexplainedWeightLoss**: Presence of unexplained weight loss, where 0 indicates No and 1 indicates Yes.
* **FatigueLevels**: Fatigue levels, ranging from 0 to 10.
* **BlurredVision**: Presence of blurred vision, where 0 indicates No and 1 indicates Yes.
* **SlowHealingSores**: Presence of slow-healing sores, where 0 indicates No and 1 indicates Yes.
* **TinglingHandsFeet**: Presence of tingling in hands or feet, where 0 indicates No and 1 indicates Yes.
* **QualityOfLifeScore**: Quality of life score, ranging from 0 to 100.

**Environmental and Occupational Exposures**

* **HeavyMetalsExposure**: Exposure to heavy metals, where 0 indicates No and 1 indicates Yes.
* **OccupationalExposureChemicals**: Occupational exposure to harmful chemicals, where 0 indicates No and 1 indicates Yes.
* **WaterQuality**: Quality of water, where 0 indicates Good and 1 indicates Poor.

**Health Behaviors**

* **MedicalCheckupsFrequency**: Frequency of medical check-ups per year, ranging from 0 to 4.
* **MedicationAdherence**: Medication adherence score, ranging from 0 to 10.
* **HealthLiteracy**: Health literacy score, ranging from 0 to 10.

**Diagnosis Information ( Target Variable)**

* **Diagnosis**: Diagnosis status for Diabetes, where 0 indicates No and 1 indicates Yes.

**Conclusion**

This dataset offers extensive insights into the factors associated with diabetes, including demographic, lifestyle, medical, and environmental variables. It is ideal for developing predictive models, conducting statistical analyses, and exploring the complex interplay of factors contributing to diabetes.

**Usability**

10.00

**License**

[Attribution 4.0 International (CC BY 4.0)](https://creativecommons.org/licenses/by/4.0/)