

Health Analytics Project: Obsessive-Compulsive Disorder (OCD) Analysis

1. Introduction

This portfolio project focuses on analyzing a dataset of **1,500 patients diagnosed with obsessive-compulsive disorder (OCD)** to uncover insights into demographics, symptom severity, treatment patterns, and trends across ethnic and socio-economic groups. The project leverages **SQL**, **Excel**, and **Power BI** for data extraction, analysis, and visualization, aiming to inform better clinical understanding and personalized treatment strategies for OCD.

2. Dataset Overview ([Click here to Download the Dataset](#))

The dataset includes the following key variables:

- **Demographics:** Age, gender, ethnicity, education level, marital status.
 - **Clinical Features:** Obsession/compulsion types, Y-BOCS (Yale-Brown Obsessive-Compulsive Scale) scores, symptom duration, medication usage.
 - **Diagnostic History:** Anxiety, depression, family history of OCD, and comorbidities.
-

3. Methodology

Tools Used

- **SQL:** Extracted and aggregated data (e.g., patient counts by ethnicity, average symptom scores).
- **Excel:** Created static visualizations (e.g., pie charts for gender distribution).
- **Power BI:** Developed an interactive dashboard for dynamic exploration of trends.

Key Queries

Examples of SQL analyses include:

1. **Gender Distribution:** 787 female vs. 553 male patients, with females having slightly higher obsession scores (18.44 vs. 17.56).
 2. **Ethnic Trends:** African patients showed rising symptom severity over time, while Asian patients exhibited stable patterns.
 3. **Symptom Severity:** Average Y-BOCS scores indicated moderate-to-severe symptoms (Obsessions: 20.05, Compulsions: 19.63).
-

4. Key Findings

Demographics

- **Age:** Average age at diagnosis: **46.63 years**.
- **Gender:** Female patients (52.5%) slightly outnumbered males (36.9%).

- **Education:** Most patients held a college degree.

Prevalence & Comorbidities

- **Anxiety:** 50% of patients had a history of anxiety.
- **Family History:** 50% reported a family history of OCD.
- **Depression:** Nearly half had no depression diagnosis, challenging assumptions about OCD-depression comorbidity.

Symptom Patterns

- **Obsession Types:** Contamination was most common (100 patients, avg. score: 19.5).
- **Compulsion Types:** Washing compulsion was prevalent (150 patients).
- **Ethnic Variations:**
 - **Hispanic patients:** Peak compulsion scores in 2017.
 - **Caucasian patients:** Spikes in symptom severity in 2018 and 2022.

Treatment & Medication

- **Medication Use:** 25.73% reported no medications; common prescriptions included SSRIs (23.47%), SNRIs (25.07%), and benzodiazepines (25.73%).
- **Symptom Duration:** Divorced individuals with college degrees experienced the longest symptom duration, while single individuals with graduate degrees had the shortest.

5. Visualizations

Power BI Dashboard (Link: [Click Here](#))

- **Interactive Elements:**
 - Line chart for monthly diagnosis trends.
 - Treemap for ethnicity distribution.
 - Bar charts for obsession/compulsion type prevalence.
- **Insights:** Highlighted seasonal diagnosis patterns (e.g., January had 20 diagnoses).

Excel Visualizations

- Static charts reinforced findings, such as gender distribution (pie chart) and ethnic trends (bar charts).

6. SQL Analysis Highlights

- **Top Obsession Types:** Contamination, symmetry, and harm-related obsessions had the highest Y-BOCS scores.
- **Ethnicity & Education:** African patients with graduate degrees had the highest obsession scores.

- **Medication Efficacy:** Patients on SSRIs had marginally lower obsession scores compared to those on benzodiazepines.
-

7. Challenges & Limitations

- **Data Gaps:** Limited information on treatment outcomes or therapy types.
 - **Temporal Trends:** Yearly symptom progression analysis was incomplete.
 - **Dataset Source:** The dataset's origin (UCI Repository) could not be verified, raising questions about reproducibility.
-

8. Recommendations

1. **Personalized Treatment:** Tailor interventions based on ethnicity, marital status, and education level.
 2. **Comorbidity Screening:** Prioritize anxiety and family history assessments during diagnosis.
 3. **Longitudinal Studies:** Track symptom progression and medication efficacy over time.
-

9. Conclusion

This project provides a robust analysis of OCD patient data, revealing critical insights into demographics, symptom severity, and treatment patterns. The integration of SQL, Excel, and Power BI enabled comprehensive exploration, while visualizations made complex trends accessible. Future work could expand on temporal trends and incorporate patient outcome data to refine treatment strategies further.
