🧠 Skin Cancer SQL Project

Advanced SQL analysis of lesion and patient data

for early-stage skin cancer detection.

🔍 Key Insights

- - Most common lesion type: ACK (461 cases)
- - Most biopsied symptoms: Itchy (326), Elevated (344), Grew (294)
- - Most biopsied demographic: Female, Pomerania background
- - Most affected body region: Face (278 lesions)
- - Largest average lesion size: MEL (14.09 mm)

🧪 Biopsy Confirmation

- Total number of biopsy-confirmed lesions: 458
- Symptoms most associated with biopsied lesions:
- - Elevated: 344
- - Itchy: 326
- - Grew: 294
- - Bleed: 195
- - Hurt: 140
- - Changed: 76

Demographic Risk Patterns

- Top Biopsied Groups:
- - Female, Non-smoking, Pesticide-exposed (Pomerania): 54 cases
- - Male, Pesticide-exposed (Germany): 42 cases
- Risk factors: Gender, Pesticide exposure, Background ethnicity

📍 Most Affected Body Regions

- Top regions with lesion presence:
- - Face: 278
- - Forearm: 219
- - Chest: 124
- - Back: 105
- Implication: High-exposure body areas more prone to lesions

📏 Lesion Size by Diagnostic Type

- - MEL: 14.09 mm (Most serious)
- - SCC: 10.46 mm
- - BCC: 9.97 mm
- - NEV and SEK are usually smaller: < 1 mm

📌 Recommendations

- - Prioritize early screening for patients with elevated, itchy, and growing lesions
- - Focus health education in regions with high face/arm exposure
- - Leverage ML-ready dataset for predictive diagnostics
- - Future studies can add UV exposure and lifestyle factors

Files in This Project

- - create_tables.sql: Table creation script
- - join_tables.sql: SQL to join tables
- - lesion_analysis.sql: Queries for insights
- - ml_ready_view.sql: Clean view for ML

✍️ Author & Tools

- Idowu Malachi
- Tools Used: PostgreSQL, pgAdmin 4
- Dataset: Skin cancer patient and lesion metadata
- LinkedIn: **linkedin.com/in/idowu-malachi-2b311a352**