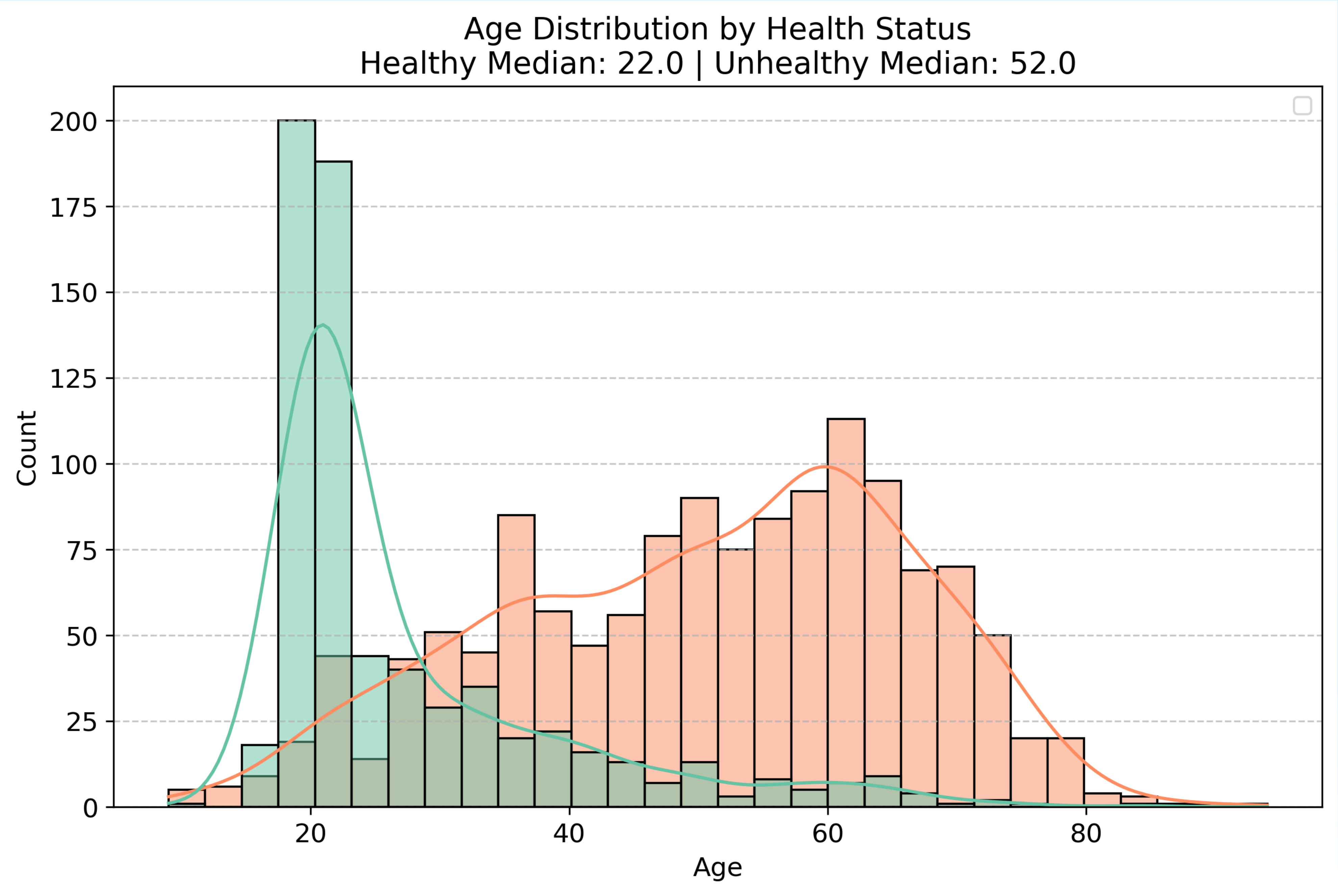
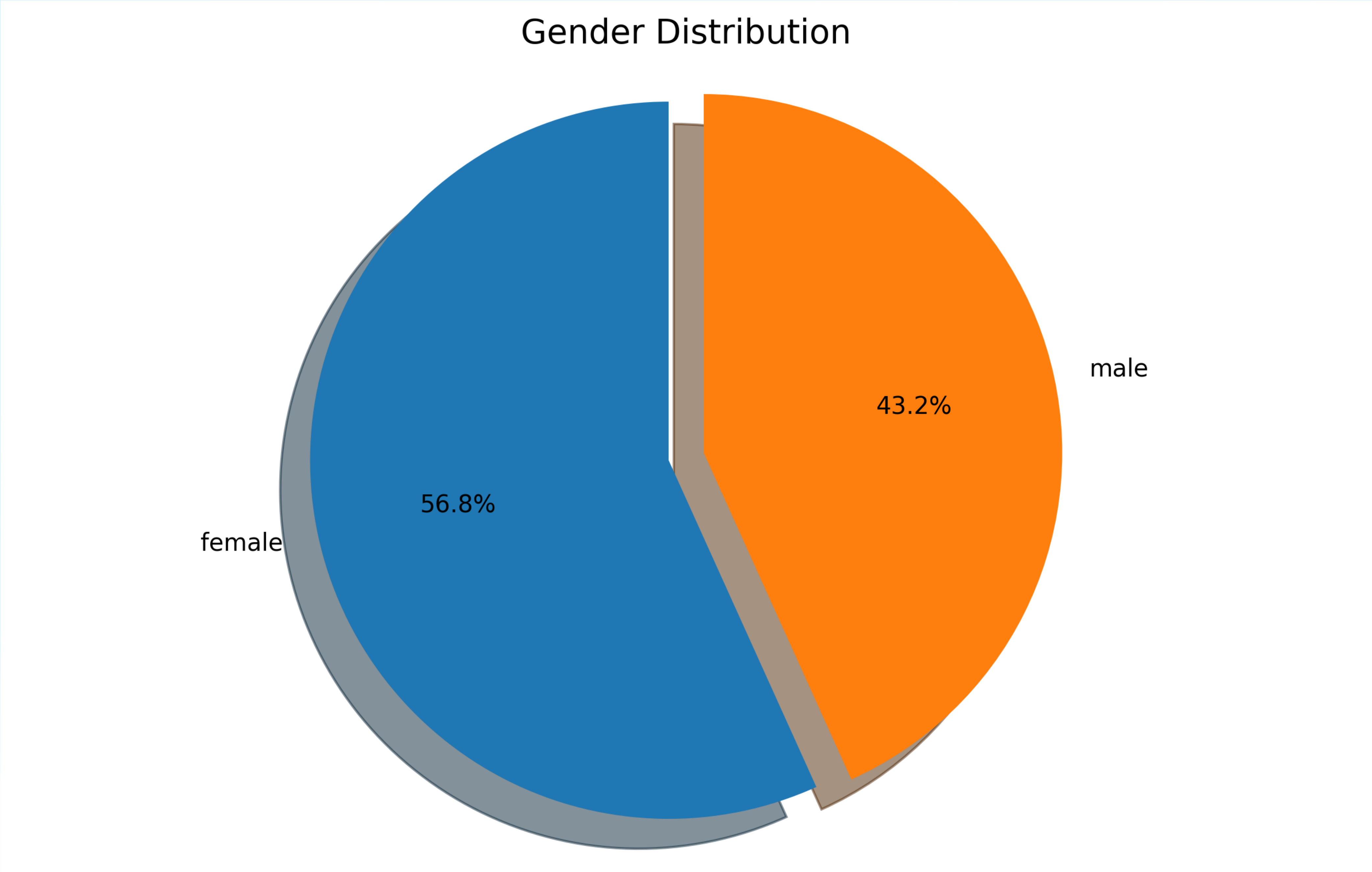
Voice Health Analysis: Predicting Health Status from Audio Features

Machine Learning Approach to Voice-Based Health Classification

Age and gender distribution

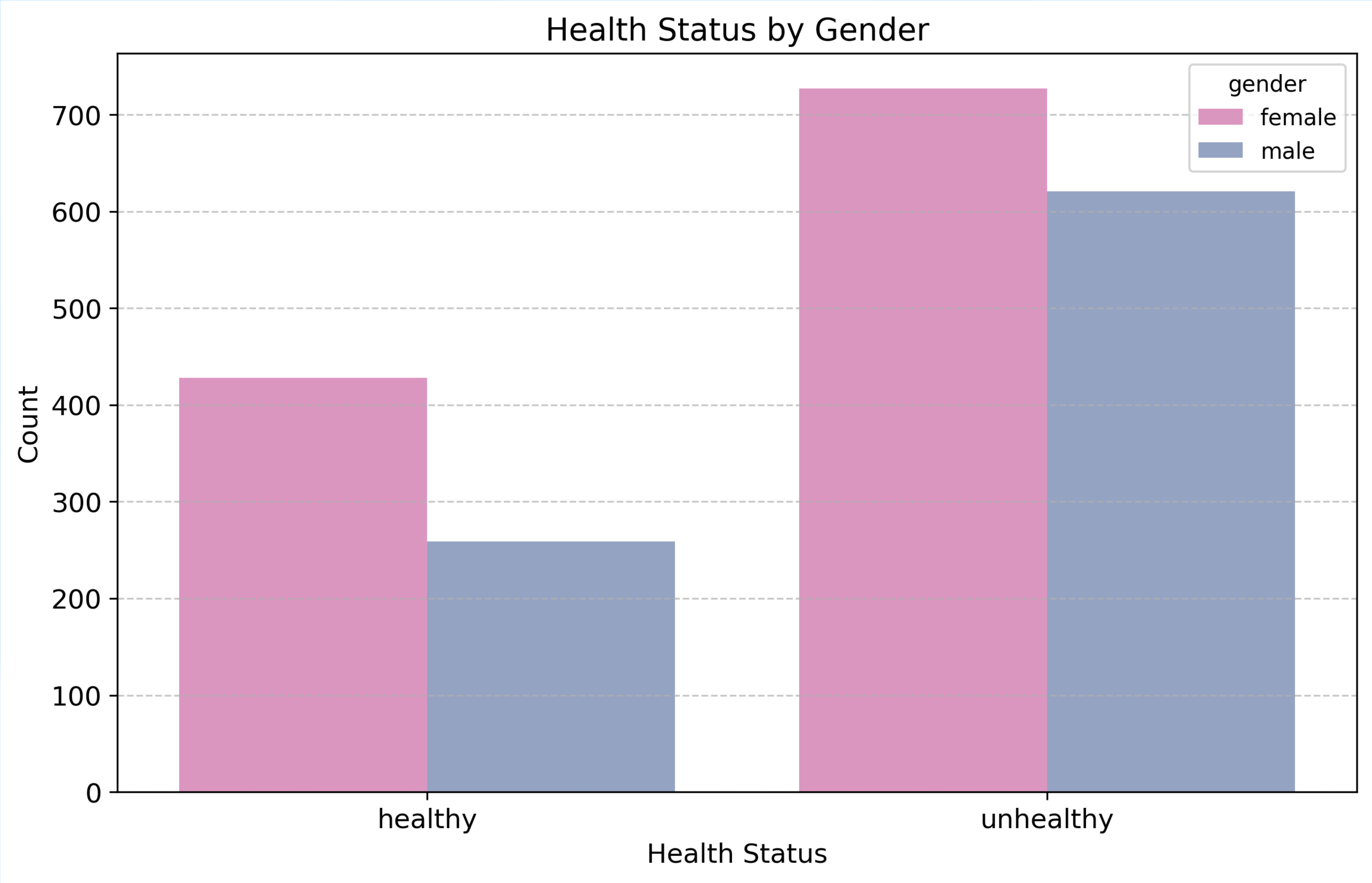




30-year age gap between healthy (median: 22) and unhealthy (median: 52) groups, presenting a significant confounding variable.

2,037 voice recordings with roughly balanced gender representation (57% female, 43% male).

Health status by gender

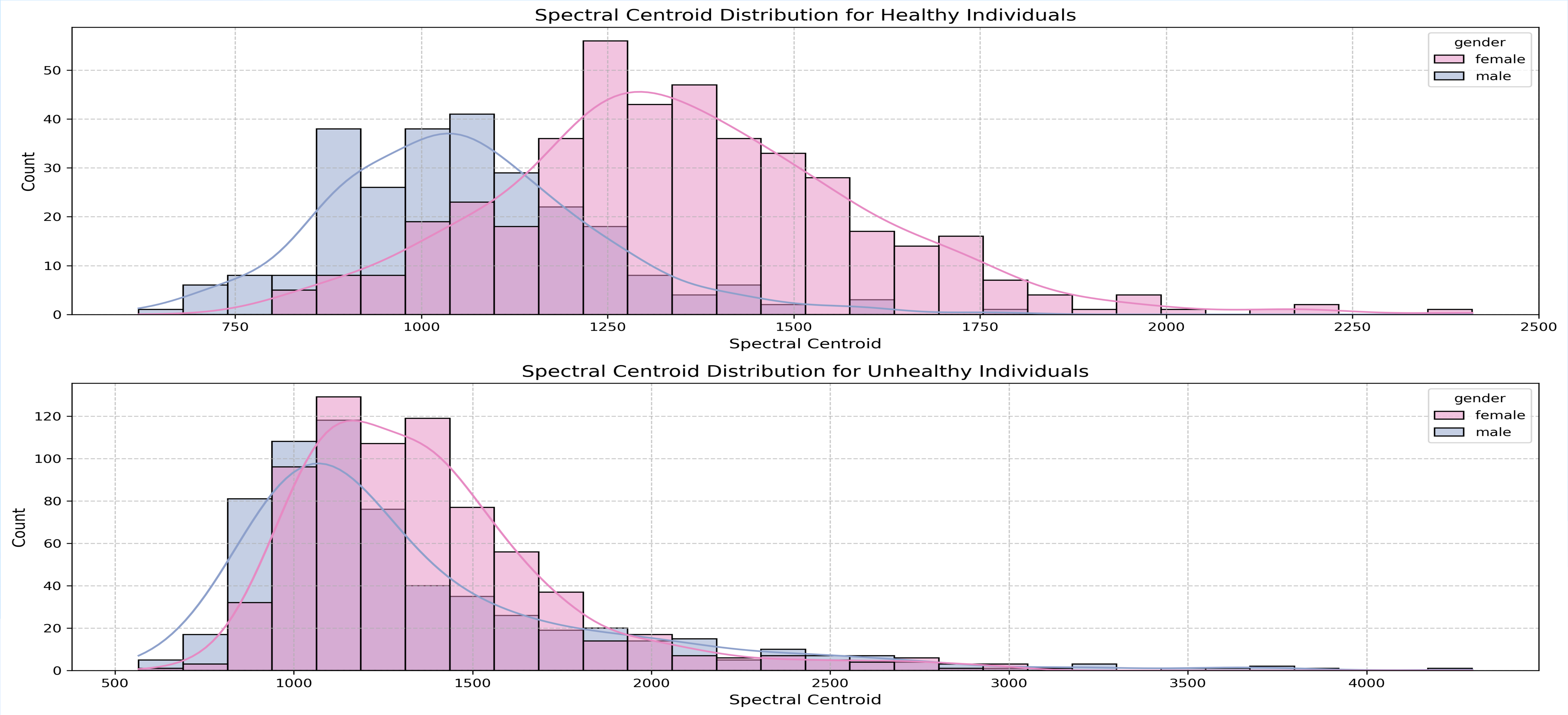


Both males and females show similar distribution patterns, with approximately twice as many unhealthy individuals compared to healthy across each gender group.

Key Discriminator: Spectral Centroid

What It Measures:

Voice brightness / center of spectral mass - think of it as the 'color' of the voice

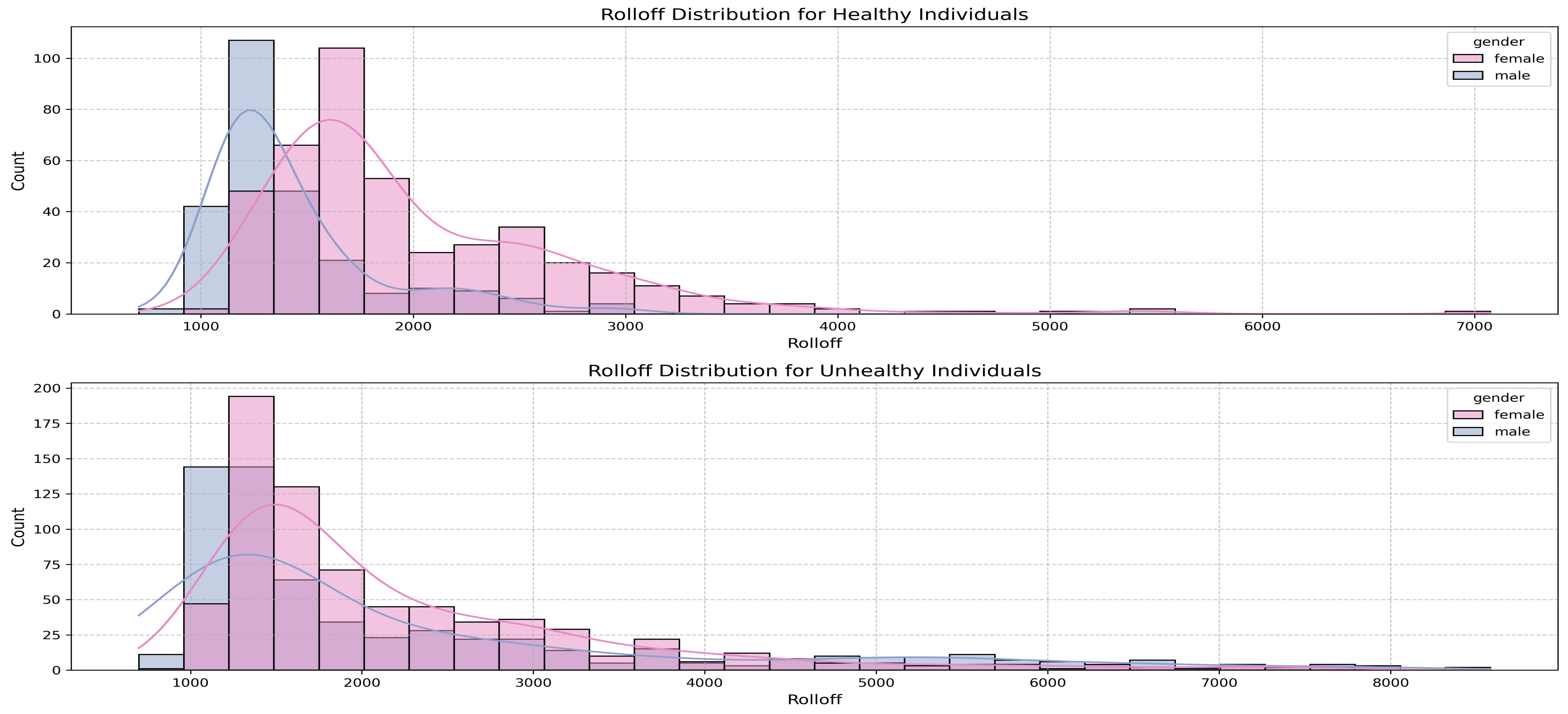


- 🔍 **Main Finding:** Unhealthy voices are significantly darker
- 📉 **Female drop:** 1250 Hz → 1100 Hz (**150 Hz decrease**)
- 📉 **Male drop:** 1050 Hz → 950 Hz (**100 Hz decrease**)
- 🎯 **Convergence effect:** Unhealthy voices become less gender-distinct

Rolloff: Indicator of Vocal Instability

What It Measures:

Frequency below which 85% of voice energy is contained - indicates breathiness and vocal stability

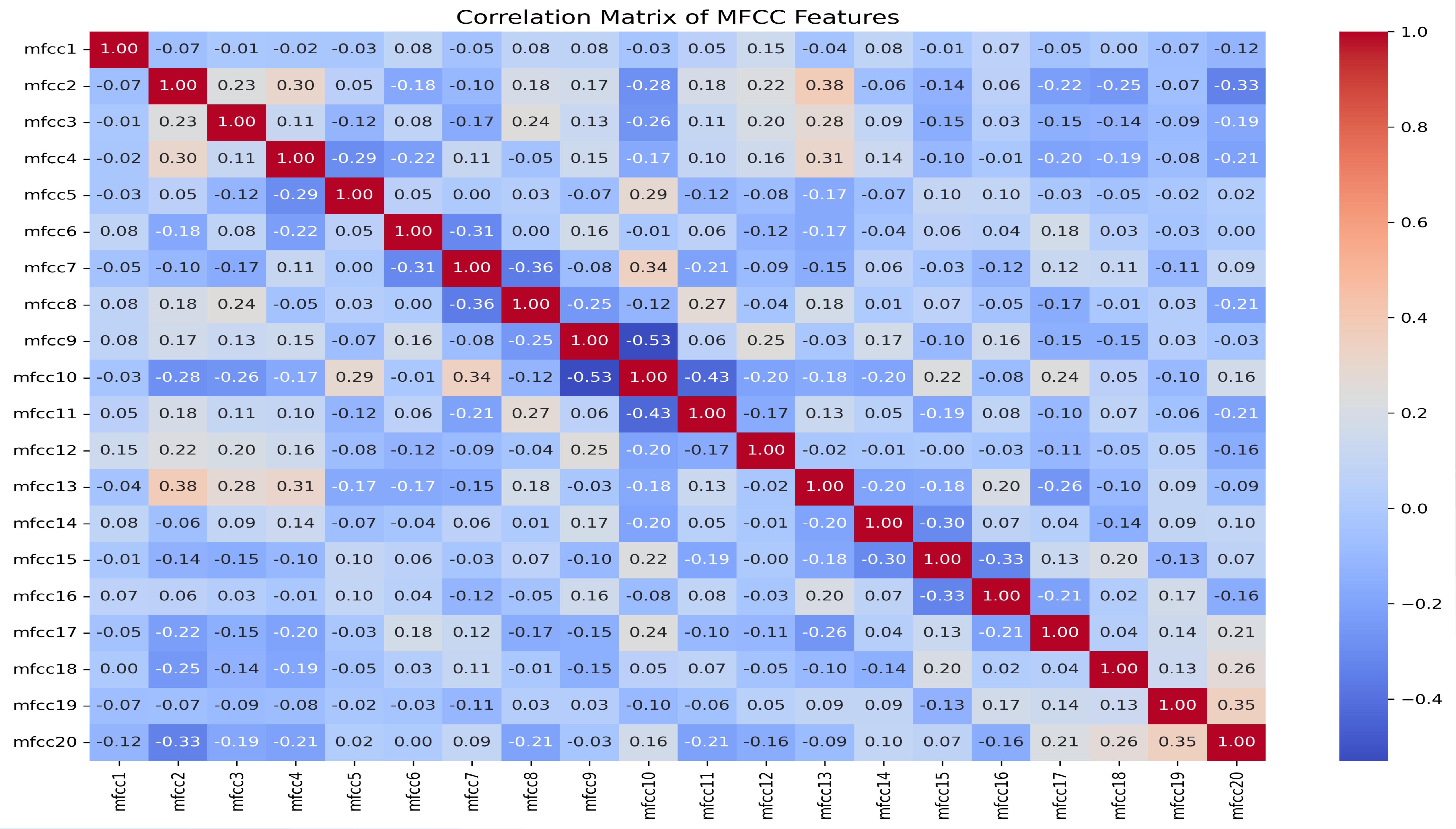


- ✅ **Healthy voices:** Tight, consistent patterns
    - Males: 1100-1300 Hz (narrow range)
    - Females: Similar stability
- ⚠️ **Unhealthy voices:** Extreme variability
    - Males: 1000-8000 Hz range (8x spread!)
    - Suggests vocal fatigue, instability, breathiness
- ℹ️ **Gender effect:** Males show dramatic changes, females stable

Feature Diversity: MFCC Correlation Matrix

What It Measures:

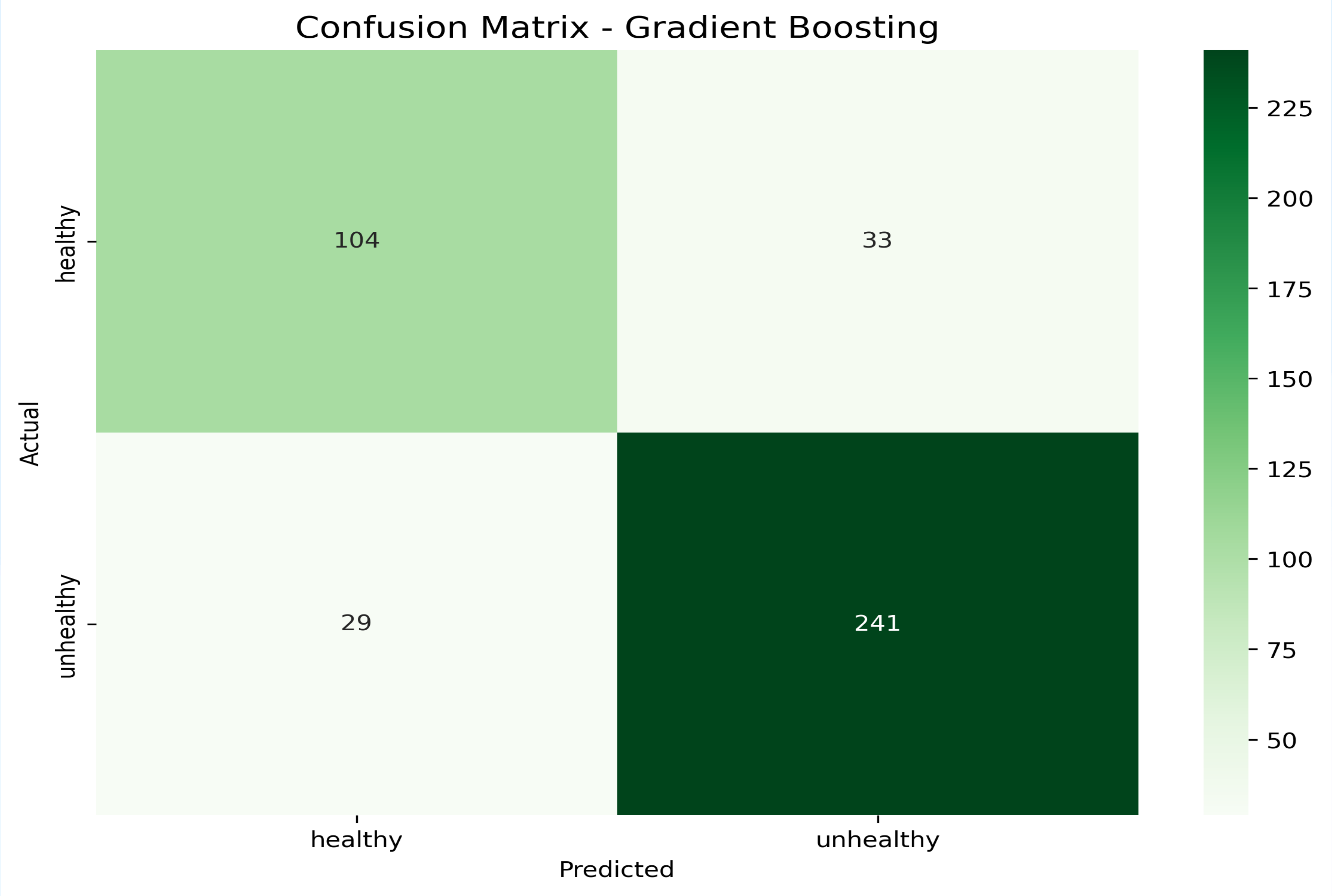
Mel-Frequency Cepstral Coefficients - capture voice characteristics the way human ears perceive sound



- 🎵 20 MFCC features (mfcc1-mfcc20)
- ✅ Low correlations (mostly < 0.4)
  - Each feature captures unique information
  - No redundancy issues
- 💡 Why it matters:
  - Diverse features = better model performance
  - No multicollinearity problems

Machine Learning Results & Key Takeaways

Best Model: Gradient Boosting



🎯 **Accuracy:** 85%

📊 **Best features: Spectral centroid, rolloff, age**

Key Findings

🎤 Voice Patterns

- Spectral centroid strongest predictor
- Rolloff shows extreme variability
- MFCCs capture nuanced differences

⚠️ Challenges

- Age confounding (30-year gap)
- Class imbalance (66/34 split)
- Gender-specific effects

🔮 Recommendations

- Age-matched validation needed
- Gender-specific models
- Control for demographic variables

Voice features show promising discriminative power, but age remains the elephant in the room