# Marketing Audience Identification - Project Plan

## 1. PREPARE Phase

### Data Collection & Organization

- Review anxiety\_attack\_dataset.csv (12,000 records)

- Create backup of raw data

- Document data dictionary and metadata

- Set up project directory structure

- Initialize version control repository

### Initial Data Assessment

- Verify data completeness for all 20 columns

- Check for missing values

- Assess data types and formats

- Document potential biases

- Identify any privacy concerns

- Note data limitations

### Tools Selection

- Primary: R for statistical analysis and visualization

- Secondary: Tableau for final visualizations

- Documentation: R Markdown for reproducible analysis

## 2. PROCESS Phase

### Data Cleaning

- Handle missing values

- Standardize occupation categories

- Clean text fields

- Remove duplicates if any

- Format dates consistently

- Document all cleaning steps

### Data Transformation

- Create derived variables:

\* Stress-severity index

\* Treatment engagement score

\* Workplace impact indicator

- Bin age groups

- Categorize occupations into professional sectors

- Create flags for high-risk indicators

#### Feature Engineering

- Create derived variables for analysis:

\* Physiological Response Index (combining HR, breathing rate, sweating)

\* Lifestyle Factor Score (combining sleep, physical activity, diet)

\* Treatment Status (combining medication and therapy)

### Quality Assurance

- Verify cleaned dataset integrity

- Cross-validate derived variables

- Document all transformations

- Create data quality report

## 3. ANALYZE Phase

### Exploratory Analysis

- Basic statistical summaries

- Distribution analysis

- Correlation analysis

- Initial segmentation exploration

### Segment Analysis

1. Primary Audience Segmentation

- Occupation & severity analysis

- Age & occupation cross-analysis

- Treatment engagement patterns

- Stress level analysis by profession

2. Behavioral Pattern Analysis

- Lifestyle factors correlation

- Treatment preferences

- Professional context impact

- Demographic patterns

3. Market Opportunity Analysis

- Segment size calculations

- Prioritization scoring

- Treatment gap analysis

- Professional context assessment

### Statistical Analysis

- Cluster analysis for segmentation

- Chi-square tests for associations

- ANOVA for group comparisons

- Regression analysis for predictors

## 4. SHARE Phase

### Data Visualization

1. Segment Profiles

- Demographic composition charts

- Attack pattern visualizations

- Treatment preference graphs

- Key pain point matrices

2. Market Opportunity

- Segment size treemap

- Opportunity matrix

- Channel preference analysis

- Journey point mapping

### Deliverables Creation

1. Technical Documentation

- Methodology report

- Statistical analysis results

- Data processing documentation

- Code repository

2. Business Deliverables

- Executive summary

- Segment profiles

- Marketing recommendations

- Channel strategy guide

3. Presentation Materials

- Executive presentation

- Detailed findings deck

- Interactive dashboard

- Handover documentation