**🔄 Complete Analysis Workflow**

flowchart TD

%% Input files

A[📁 abundance\_data.csv] --> M

B[📁 metadata.tsv] --> M

%% Main pipeline decision

M{🤔 Choose Analysis Approach}

%% Main automated pipeline

M -->|Automated Pipeline| N[🚀 main\_pws.py]

%% Manual pipeline

M -->|Manual Pipeline| P1[📊 pws.py]

%% Main pipeline execution

N --> N1[📊 pws.py]

N1 --> N2[🔬 filtering5.py]

N2 --> N3[📈 plotting\_script.py]

N3 --> N4[📦 Organize in out\_pwd/]

%% Manual pipeline execution

P1 --> P2[🔬 filtering5.py]

P2 --> P3[📈 plotting\_script.py]

%% Intermediate files

N1 --> F1[📄 parameter\_weighted\_analysis.csv]

N1 --> F2[📄 top30.txt]

P1 --> F1

P1 --> F2

N2 --> F3[📄 filtrado.csv]

N2 --> F4[📄 filtrado\_conservativetest.csv]

N2 --> F5[📄 sup\_statistics.csv]

N2 --> F6[📄 filtering\_report.txt]

P2 --> F3

P2 --> F4

P2 --> F5

P2 --> F6

N3 --> F7[🖼️ Parameter\_weighted\_barplot.png]

N3 --> F8[🖼️ Parameter\_weighted\_barplot2.png]

P3 --> F7

P3 --> F8

%% Decision point for Welch's alternative

F1 --> D{🧪 Need Welch's t-test?}

F2 --> D

%% Welch's alternative pathway

D -->|Yes - Conservative Analysis| W1[🔬 filtering\_welch.py]

D -->|No - Analysis Complete| END1[✅ Standard Analysis Complete]

W1 --> W2[📄 filtrado1.csv]

W1 --> W3[📄 welch\_filtering\_report.txt]

W2 --> W4[📈 plot\_weighted\_scores.py]

W4 --> W5[🖼️ parameter\_weighted\_barplot\_welch.png]

W5 --> END2[✅ Welch's Analysis Complete]

%% Styling

classDef inputFile fill:#e1f5fe

classDef script fill:#f3e5f5

classDef output fill:#e8f5e8

classDef decision fill:#fff3e0

classDef endpoint fill:#ffebee

class A,B inputFile

class N,N1,N2,N3,P1,P2,P3,W1,W4 script

class F1,F2,F3,F4,F5,F6,F7,F8,W2,W3,W5 output

class M,D decision

class END1,END2 endpoint

**📋 Detailed Workflow Steps**

**🎯 Option 1: Automated Pipeline (Recommended)**

📁 Input Files

├── abundance\_data.csv

└── metadata.tsv

↓

🚀 main\_pws.py (Orchestrator)

├── 📊 pws.py

├── 🔬 filtering5.py

├── 📈 plotting\_script.py

└── 📦 out\_pwd/ organization

↓

📂 Complete Results in out\_pwd/

├── 📄 parameter\_weighted\_analysis.csv

├── 📄 top30.txt

├── 📄 filtrado.csv (individual tests)

├── 📄 filtrado\_conservativetest.csv (conservative)

├── 📄 sup\_statistics.csv (detailed stats)

├── 📄 filtering\_report.txt

├── 🖼️ Parameter\_weighted\_barplot.png

├── 🖼️ Parameter\_weighted\_barplot2.png

└── 📄 pipeline\_log.txt

**🔧 Option 2: Manual Pipeline**

📁 Input Files → 📊 pws.py → 🔬 filtering5.py → 📈 plotting\_script.py

↓ ↓ ↓ ↓

Core Analysis Dual Filtering Comparative Plots

**🧪 Option 3: Welch's t-test Alternative**

📄 parameter\_weighted\_analysis.csv ──┐

📄 top30.txt ──┤

↓

🔬 filtering\_welch.py

↓

📄 filtrado1.csv

📄 welch\_filtering\_report.txt

↓

📈 plot\_weighted\_scores.py

↓

🖼️ parameter\_weighted\_barplot\_welch.png

**🎮 Command Sequences**

**Full Automated Analysis:**

python main\_pws.py

# ✅ Complete analysis with organized outputs

**Manual Step-by-Step:**

python pws.py # Core PWS analysis

python filtering5.py # Hybrid statistical filtering

python plotting\_script.py # Comparative visualizations

**Add Welch's Alternative:**

# After main analysis is complete:

python filtering\_welch.py # Welch's t-test filtering

python plot\_weighted\_scores.py # Welch-specific visualization

**Complete Comparative Analysis:**

python main\_pws.py # Full pipeline

python filtering\_welch.py # Add Welch's alternative

python plot\_weighted\_scores.py # Welch's visualization

# Result: All three approaches available for comparison

**📊 Output File Summary**

| **Stage** | **Individual Tests** | **Conservative Tests** | **Welch's t-test** |
| --- | --- | --- | --- |
| **Filtering** | filtrado.csv | filtrado\_conservativetest.csv | filtrado1.csv |
| **Visualization** | \*\_barplot.png | \*\_barplot2.png | \*\_barplot\_welch.png |
| **Statistics** | sup\_statistics.csv | filtering\_report.txt | welch\_filtering\_report.txt |

**🔄 Decision Tree**

Start Analysis

↓

Need automated workflow?

├── Yes → main\_pws.py → Complete Results

└── No → Manual scripts → Intermediate Results

↓

Need conservative approach?

├── Yes → Already included in main results

└── No → Continue

↓

Need standardized Welch's test?

├── Yes → filtering\_welch.py + plot\_weighted\_scores.py

└── No → Analysis complete

**💡 Recommendations by Use Case:**

* 🔬 **Research & Exploration**: main\_pws.py (hybrid approach)
* 📚 **Academic Publication**: main\_pws.py + Welch's alternative
* 🏛️ **Regulatory Submission**: Welch's alternative only
* 🔍 **Method Comparison**: All three approaches
* ⚡ **Quick Analysis**: main\_pws.py (automated)

**⚠️ Dependencies:**

* Welch's scripts **require** parameter\_weighted\_analysis.csv and top30.txt
* Generate these first using pws.py or main\_pws.py
* All approaches use the same input files (abundance\_data.csv, metadata.tsv)