Proof of Taste Project Plan

Trey Wood

July 23, 2025

Project Overview

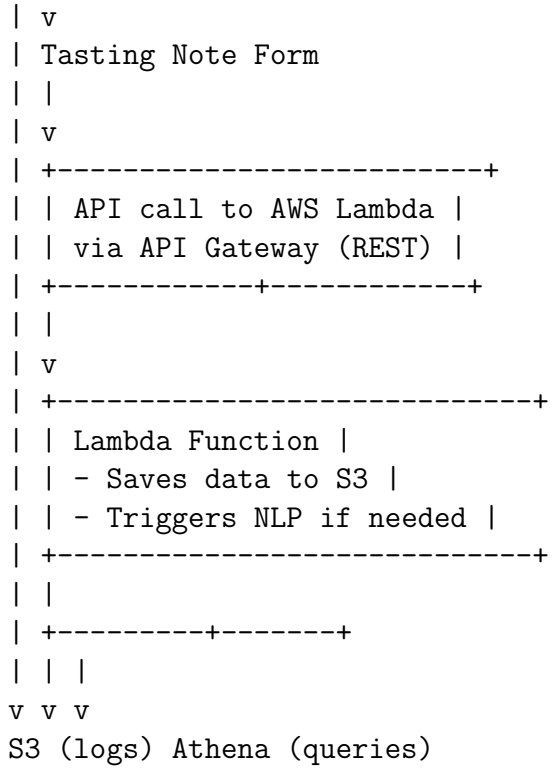
Goal: Build a bourbon-tasting tracker and recommendation engine using:

- GitHub for code management and version control
- Streamlit Cloud for a public-facing web app
- AWS (S3, Lambda, Athena) for data storage and cloud orchestration

App Features:

- User selects or inputs a Subject ID
- Users submit bourbon tasting notes via a Streamlit app
- Notes are stored in AWS S3, and optionally analyzed with NLP
- A recommender suggests bourbons based on taste profiles
- Athena provides SQL-based analytics (optional)

Architecture Overview



Phased Work Plan

Phase 1: GitHub + Local Streamlit MVP (Week 1)

Goal: Basic local app with Subject ID and tasting form.

- Create GitHub repo: `proof-of-taste`
- Set up Streamlit locally
- Build form UI for submitting and viewing notes
- Save data to local JSON files (e.g. `data/{subject_id}.json`)

Phase 2: Streamlit Cloud Hosting (Week 2)

Goal: Deploy public-facing app

- Push code to GitHub
- Create `requirements.txt`, `secrets.toml`
- Deploy via Streamlit Cloud

Phase 3: AWS S3 Integration (Week 3)

Goal: Move tasting logs to cloud storage

- Set up S3 bucket (`proof-of-taste-data`)
- Use `boto3` to save/load JSON logs
- Secure access using IAM roles and keys

Phase 4: AWS Lambda API (Week 4)

Goal: Add REST API via Lambda + API Gateway

- Build ingestion Lambda function
- Connect via API Gateway
- Update Streamlit to POST form data to API

Phase 5: NLP + Recommendations (Week 5–6)

Goal: Add flavor NLP and suggestion engine

- Use spaCy to extract flavors from notes
- Track preferences per Subject ID
- Recommend bourbons from curated dataset

Phase 6: Athena Query Layer (Optional)

Goal: Enable analytics and ad hoc queries

- Enable Athena over S3 logs
- Create SQL tables/views
- Expose insights via Streamlit dashboard

Repository Structure

```
proof-of-taste/
|- app.py # Streamlit UI
|- requirements.txt
|- utils/
| |- s3_utils.py # AWS S3 helper
| |- nlp_utils.py # Flavor extraction
| |- recommender.py # Matching logic
```

```
|- data/
| |- bourbon_db.csv # Bourbon metadata
|- .streamlit/
|- secrets.toml # AWS keys (not committed)
```

Tool Summary

Layer	Tool	Purpose
UI	Streamlit	Front-end app (deployed to Streamlit Cloud)
Code Repo	GitHub	Version control and CI/CD trigger
Storage	AWS S3	Log storage and raw data
Compute	AWS Lambda	Ingestion and analysis endpoint
SQL Query	AWS Athena	Optional analytics
Auth	Subject ID input	Light ID management (expandable)

Prepared for personal project development and AWS portfolio demonstration.