**Sprint 1: Data Exploration & Basic Functionality**

**1) User can see homepage of the app---All Good**

* **Subtasks**
  + **UI Design**: Create a welcoming homepage layout summarizing key ESG categories (S, E, G) and six main sub-categories (Social/Governance Risks/Opportunities, etc.).--JC
  + **Navigation**: Implement buttons/links leading users to select industry/company or view general stats.-PS
  + **Basic Data Summary**: Display a few highlighted stats (e.g., total number of companies in the dataset, date range).-VN
  + **Testing**: Validate layout and quick data fetch on multiple browsers. -VP

**2) User can log in -------All Good**

* **Subtasks**
  + **Frontend Login Form**: Company, Industry account based login with basic authentication. ---JC
  + **Session Management**: Ensure user remains logged in across pages.--PS
  + **Unit Tests**: Validate correct login, logout, and error handling.--VN

**3) User can view descriptions of data metrics-------All Good**

* **Subtasks**
  + **Backend Endpoint**: Return metric name, category, unit, and descriptions from the existing ESG dataset.---SB
  + **Frontend Catalog**: Create a UI listing each metric under its respective category (Env. Risk, Env. Opportunity, etc.).---PS
  + **Search/Filter**: Let users type a keyword or category name to narrow metric list.---VN
  + **Verification**: Ensure these metadata match the Clarity AI definitions.--VP

**4) User can explore basic stats of the data-------All Good**

* **Subtasks**
  + **Stat Computations**: Calculate mean, median, standard deviation, etc., by metric.--VP
  + **Simple Visuals**: Bar charts or tables to show aggregated data (all companies combined) for each metric.---VN
  + **Time Range Toggle**: Ability to pick a specific year or multiple years (from 2016–2024) for the stats.--JC
  + **QA Testing**: Cross-check a sample subset of data calculations for correctness.—PS,SB

**5) User can select an industry or specific company-------All Good**

* **Subtasks**
  + **Frontend Dropdowns**: Retrieve industry and company lists from backend.---JC
  + **Filtering Logic**: Once selected, only those relevant records show in subsequent stats and visuals.--VN
  + **Comparison Option**: (If feasible) letting the user compare multiple companies from the same industry. VP,SB
  + **Integration Test**: Confirm the correct data is fetched per user choice.PS

**6) User can view and select metrics from each category-------All Good**

* **Subtasks**
  + **UI for Metric Selection**: Checkboxes or multi-select to choose from Environmental, Social, Governance metrics.--PS
  + **Backend Endpoints**: Filter dataset based on user‐selected metrics only.—VP,VN
  + **Show Filtered Results**: Display tables/graphs focusing only on selected metrics (e.g., Social Risk vs. Social Opportunity).—PS,SB
  + **Testing**: Validate that queries match the chosen metrics accurately.JC

**7) User can explore time‐series trends (2016–2024) -------All Good**

* **Subtasks**
  + **Year‐by‐Year Aggregation**: Summarize data for each year in the chosen range.---VP
  + **Line/Bar Charts**: Visualize changes over time for selected metrics (e.g., CO2 Direct Scope 1 from 2016–2024).---PS
  + **Frontend Interactivity**: Sliders or dropdowns to specify the year range.--VN
  + **Data Validation**: Check for missing data in certain years; handle gracefully in visualizations.—JC,SB

**Sprint 2: Advanced EDA & PCA (Ontology Integration)**

**1) User can incorporate industry-matched data (Semiconductors & Pharmaceuticals) into EDA---TBD put in sprint 1**

**Subtasks:**

* **Integrate SASB Industry Files:** Pull the pre-uploaded files (Semiconductors, Pharma) from group folder into the existing data pipeline.--SB
* **Automated Matching Script:** Develop or refine a script that cross-references raw ESG data with these industry classification files for better alignment in the EDA.--VP
* **Include in EDA:** Show new EDA views for selected industries, highlighting coverage (e.g., # of companies matched vs. unmatched).—PS,JC
* **Testing:** Validate that matching process is accurate (e.g., correct industry codes, no data duplication).---VN

**2) User can view an aggregated ESG metric for each of the six categories -------All Good**

**Subtasks**

* **Backend Aggregation Logic:** Implement the concept of a single ‘representative’ metric per category (e.g., “ERiskMetric” for Environmental Risk).---VP
* **UI for Category Setup:** Let the user choose which underlying metrics from the dataset feed into the aggregated metric (or show a default set).
* **Model Definition:** Provide placeholders or formulas (e.g., weighting factors) that define how sub-metrics combine into the aggregated metric.SB,PS,JC
* **Tests:** Confirm that the final aggregated metric values (for each category) match expectations in a few sample companies--VN

**3) User can select a model & click “Calculate” to compute each category’s aggregated metric. -------All Good**

* **Subtasks**
  1. **Frontend Workflow**:
     + Step 1: User picks a metric for each ESG category.--JC
     + Step 2: User selects (or confirms) a *model/method* for combining sub metrics.--PS
     + Step 3: Click **“Calculate”** to generate final aggregated values.—VN,SB
  2. **Metrics Computing Layer**:
     + Integrate user selections into the calculation pipeline.
     + Perform relevant computations (weighting, normalization, sums, etc.).
  3. **Real-time Feedback**: NO GO
     + Display the aggregated metric result for each category in a UI panel.
     + Optionally highlight anomalies (e.g., missing sub-metrics).
  4. **Validation**:
     + Thoroughly test a sample company or industry to ensure calculation correctness. ---VN,JC

**4) User can perform standard PCA on aggregated ESG category metrics. -------All Good**

* **Subtasks**
  1. **Backend PCA Module**:
     + Extend or reuse existing PCA code to run on the 6 aggregated category metrics, for chosen industries (Semiconductors, Pharma. ---🡪SB
  2. **Normalization**:
     + Automatically standardize the 6 aggregated metrics prior to PCA. ----SB
  3. **Scree & Biplot**:
     + Visualize the explained variance by each principal component.
     + Show biplot to interpret aggregated categories’ influence. ---------🡪VP, JC
  4. **Testing**:
     + Compare results to manually computed examples or smaller test sets. –VN, PS

**5) User can run ontology-enhanced PCA and compare with standard PCA ---All Good**

* **Subtasks**
  1. **Ontology-Aware Weighting**:
     + If a user selects “Ontology-Enhanced PCA,” retrieve relevant ontology weighting or relationships from the knowledge-graph microservice (e.g., some metrics might have higher weights if flagged “material” in that industry).----🡪VP,SB
  2. **Computational Pipeline**:
     + Integrate those ontology-based weights into the same PCA pipeline. ---SB,JC
  3. **Comparison View**:
     + Show side-by-side difference in the explained variance or principal components vs. the standard PCA.----VN,JC
  4. **Integration Testing**:
     + Confirm data flow from ontology microservice to metrics computing layer is correct. ---VP

**Sprint 3: Final Integration, Advanced Refinements, & Wrap-Up-🡪---All Good**

**1) User can extend coverage to additional industries beyond Semiconductors & Pharmaceuticals-**

(If time permits, or keep it minimal but feasible)

Subtasks

* **SASB/IFRS Industry Extension:** Integrate more industry-matched files or refine existing script to handle multiple new industries.---VP,VN
* **UI Industry Selector:** Update the GUI to include any newly added industries in a dropdown/filter.---JC,PS
* **Verification:** Quickly test the extended EDA pipeline (aggregated metrics, PCA) for at least one new industry (e.g., “Automobile” or “Energy”).---SB
* **Documentation:** Note each newly supported industry in the system help or user guide.---SB

**2) User can detect anomalies or outliers in ESG data (Optional advanced feature)**

* **Subtasks**
  1. **PCA Reconstruction Error**: Implement a function that uses PCA to detect potential outliers in aggregated ESG category metrics.
  2. **UI Indicators**: Highlight or tag data points with suspiciously high reconstruction error or that deviate significantly from industry norms.
  3. **Exploratory Flag**: Provide an “anomaly list” for Semiconductors and Pharma to demonstrate the feature.
  4. **Testing**: Validate approach with a known subset or artificially injected outliers.

**3) User can download final computed metrics & results-🡪---All Good**

* **Subtasks**
  1. **Export Options**: Provide CSV, Excel, or PDF export of final aggregated metrics, PCA loadings, or even anomaly flags using Tableu----JC,VN
  2. **Batch Export**: Allow user to export data across multiple companies or industries at once.
  3. **Testing**: Confirm correct file formatting (no missing columns, consistent naming).---VN

**4) User can interact with a final integrated dashboard-🡪---All Good**

* **Subtasks**
  1. **Consolidated UI**: Combine the EDA, aggregated metric calculator, PCA (both standard & ontology-enhanced), anomaly detection, and final visualizations into a single, coherent dashboard.---PS,VN
  2. **Refined Navigation**: Ensure seamless flow from industry selection → EDA → metric modeling → PCA → anomalies → final summary.-🡪SB,PS,JC,VP,VN
  3. **Performance Tuning**: Optimize front-end load times and backend queries for a smoother user experience.-🡪 SB,PS,JC,VP,VN
  4. **Cross-Browser & Cross-Device Testing**: Confirm that the integrated dashboard works well in standard desktop browsers and (if feasible) mobile/responsive displays.

**5) Project Wrap-Up & Documentation-🡪---All Good**

* **Subtasks**
  1. **Final System Documentation**: Summarize the pipeline, data management design, knowledge graph usage, and model computations.SB,VP,JC,VN,PS
  2. **Technical Manual**: Outline steps to maintain or extend the system (e.g., how to add new industries, new aggregated metrics, or new computations).-🡪 SB,PS,JC,VP,VN
  3. **User Guide**: Provide instructions for typical usage scenarios (e.g., how to select an industry, run PCA, interpret results).-🡪 SB,PS,JC,VP,VN
  4. **Final Demo Preparation**: Prepare a stable build and demo script for the final presentation.-🡪 SB,PS,JC,VP,VN