GRADUATE CERTIFICATE

INTELLIGENT REASONING SYSTEM (IRS)

-- **PROJECT REPORT** --

BiasTrack: Correcting Gender Pay Gaps Before They Widen

GROUP 9

Arshi Saxena (A0331999J)

Pranjali Rajendra Sonawane (A0326167B)

Table of Contents

1 Introduction

Pay equity remains a critical organizational and ethical goal in today's business landscape, where companies are increasingly expected to uphold fairness, inclusivity, and transparency in their compensation structures. Despite progressive awareness and diversity initiatives, gender-based pay disparities continue to exist across industries and roles. These inequities, often subtle and systemic, can affect employee morale, talent retention, and organizational reputation, making it essential for businesses to proactively identify and address them.

Traditional compensation assessments are typically manual, relying on broad benchmarking or static reports that offer limited visibility into the root causes of pay gaps. Human Resource (HR) professionals often find themselves equipped with data, but without the analytical clarity or tools needed to translate that data into fair and actionable pay decisions. As organizations evolve towards evidence-based HR practices, there is an urgent need for intelligent systems that go beyond descriptive reporting—systems capable of uncovering hidden biases, quantifying disparities, and recommending equitable solutions that balance fairness with financial practicality.

BiasTrack is developed as a Minimum Viable Product (MVP) that addresses this precise need. It is designed as an intelligent reasoning and optimization system that empowers HR teams to analyse, understand, and act upon pay disparities in a transparent and data-driven manner. The system integrates machine learning, regression analysis, and interactive visualization into a unified platform that bridges the gap between analytics and decision-making.

At its core, BiasTrack uses regression techniques to model salary outcomes based on multiple influencing factors such as gender, education, performance evaluation, and seniority. Through these models, it identifies how each attribute contributes to overall compensation patterns and predicts what a fair salary distribution should look like. The insights are then visualized through user-friendly dashboards that make complex statistical findings accessible and interpretable for HR professionals.

By combining predictive modelling with interpretability and interactivity, BiasTrack transforms complex data into clear insights that inform strategic decision-making. The system's strength lies not only in its analytical capability but also in its design philosophy—prioritizing fairness, explainability, and usability. This approach ensures that HR professionals can make data-backed compensation decisions that are equitable, defensible, and aligned with organizational values.

Ultimately, BiasTrack represents a step toward bridging technology and human judgment in pursuit of workplace equity. It demonstrates how intelligent systems can play a transformative role in advancing fair pay practices, equipping organizations with the ability to monitor, understand, and proactively **correct gender-based pay gaps before they widen**.

2. Market Context

2.1. Problem

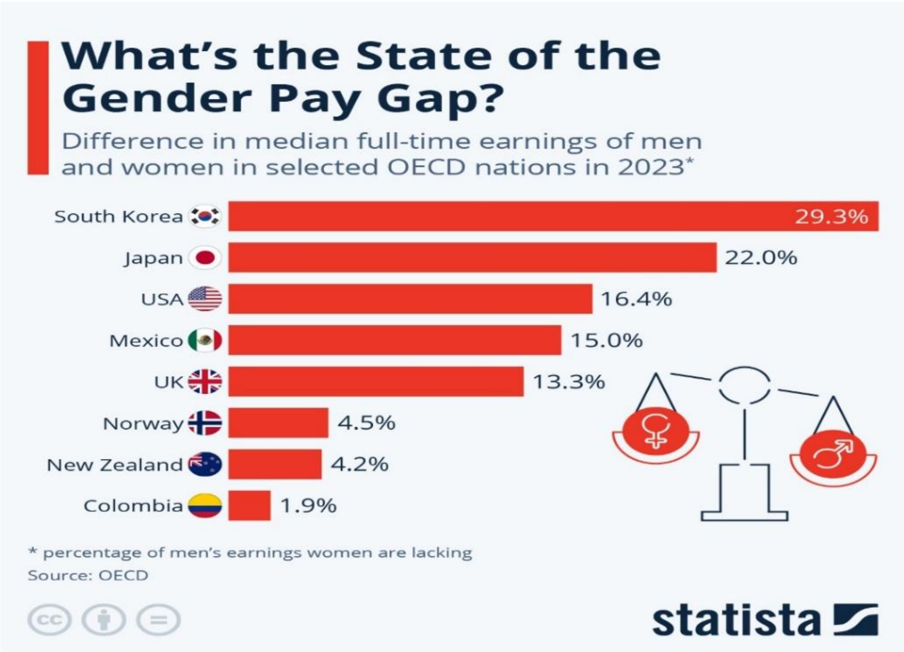
Gender-based pay disparities remain a significant challenge for organisations globally and locally, with material consequences for fairness, talent management and organisational reputation.

According to the World Economic Forum's Global Gender Gap Report 2025, the gender gap across 148 economies is only **68.8% closed**, meaning that on average women still face a significant disparity in income and career outcomes. In many markets, women earn approximately **83 cents for every dollar earned by men**, translating to a gap of roughly 17 percent.

In the Singapore context, the latest data from the Ministry of Manpower shows that the unadjusted median gender pay gap for full-time employees aged 25-54 declined from **16.3 % in 2018 to 14.3 % in 2023**. When adjusted for human-capital and labour-market factors the gap narrows further to about **6.0 %**.

These numbers reveal two important observations: first, the raw gap remains sizeable; second, even when controlling for factors such as education, seniority and role, a non-negligible 'unexplained' gap persists — pointing to structural and systemic issues rather than purely measurable differences in inputs.

At an organisational level, the implications are compelling. Unaddressed pay gaps can erode employee trust, hinder diversity and inclusion goals, and expose businesses to regulatory or reputational risks — especially as pay transparency becomes more prominent in stakeholder expectations. Against this backdrop, companies require tools that can not only **detect** disparities but also **explain** and **address** them in a financially viable manner.



2.2. Business Opportunity

The growing regulatory, societal and investor focus on pay equity creates a clear market demand for systems that deliver actionable insight into gender pay disparities. For example, transparency laws in multiple jurisdictions—such as the **EU Pay Transparency Directive**—now require companies to disclose gender pay gap data and salary-range information. While Singapore does not yet mandate full pay reporting, evolving guidelines under the Workplace Fairness Act and other frameworks signal an emerging need for automated tools.

From a commercial standpoint, HR analytics solutions positioned to offer transparency, simulation, and optimisation capabilities are increasingly valuable. Organisations see opportunity in leveraging **regression-based analytics** and **interactive dashboards** to anticipate pay-gap liabilities, model adjustment scenarios, and align compensation practices with fairness and business constraints. This demand positions a system like BiasTrack—which integrates detection, explanation and simulation—as a timely and strategic solution.

Moreover, by offering a user-friendly interface for HR decision-makers rather than purely presenting dashboards, BiasTrack addresses a gap not well met by traditional compensation benchmarking tools. In doing so, it provides a differentiated value proposition: bridging analytics with actionable decision-support, underpinned by machine learning and visualisation.

In summary, the confluence of persistent pay gaps, evolving regulatory expectations and the demand for actionable HR intelligence creates a strong context and opportunity for the project.

3. Market Research

3.1. Industry Overview

The global HR analytics and compensation intelligence market has grown rapidly over the past decade, driven by the increasing emphasis on pay transparency, diversity, and evidence-based workforce management. According to *Grand View Research (2025)*, the HR analytics market is valued at **USD 4.8 billion** and is projected to grow at a **compound annual growth rate (CAGR) of 14 %** through 2030.

Pay-equity analytics, a subset of this market, is emerging as one of the fastest-growing segments as organisations face rising regulatory and social pressure to disclose and rectify wage disparities.

Businesses are increasingly shifting from static, compliance-oriented reporting toward **decision-support tools** that integrate predictive analytics and simulation. As AI and machine learning become integral to HR functions, companies are seeking solutions that combine *explainability*, *interactivity*, and *financial feasibility*—precisely the niche that BiasTrack addresses.

3.2. Key Competitors & Offerings

The current market is served by a range of HR technology platforms that focus on pay-equity analysis, benchmarking, and compliance. Some of the key players include:

Company	Core Focus	Key Differentiators / Limitations
Syndio	Pay-equity analytics and compliance reporting	Strong compliance and audit features, but limited interactivity or optimization; mainly U.S. focused.
Ravio	Real-time compensation benchmarking across industries	Offers salary insights and peer comparison, but lacks transparency on algorithmic reasoning.
Gapsquare	Data analytics platform for pay-gap reporting	Provides clear dashboards, but focuses primarily on detection rather than simulation or recommendations.
Payscale	Salary benchmarking and market data platform	Offers extensive salary data but not designed for organizational pay-equity correction.

Existing solutions primarily highlight *where* pay inequities exist but rarely explain *why* they occur or *how* organizations can resolve them within financial and hierarchical constraints. This gap opens the door for **BiasTrack's differentiator — transparent regression-based insights coupled with real-time simulation and optimization.**

3.3. Market Trends

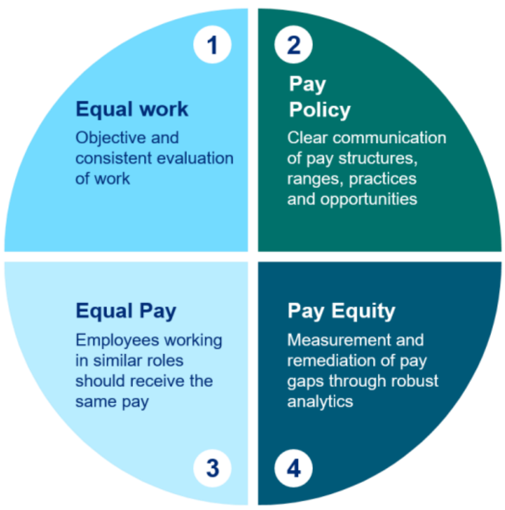
- **Regulatory Push for Transparency:**
  New laws such as the **EU Pay Transparency Directive (2026)** and similar U.S. state-level pay-equity acts are accelerating corporate investment in analytics tools capable of generating compliant reports and corrective recommendations.
- **Integration of Explainable AI in HR:**
  HR teams increasingly demand explainable models rather than black-box predictions to maintain employee trust and meet ethical-AI standards.
- **Rise of Predictive and Prescriptive HR:**
  Companies are moving from descriptive dashboards to *predictive* ("what will happen") and *prescriptive* ("what should we do") analytics.
- **Budget-Constrained Optimization:**
  With tightening corporate budgets, decision-makers prefer tools that factor in constraints — aligning fairness improvements with financial sustainability.

These trends validate the relevance of BiasTrack's approach: combining regression-based prediction with interpretability and constraint-based optimization, delivered through a user-friendly interface.

3.4. Opportunity for BiasTrack

BiasTrack occupies a strategic position at the intersection of **AI-driven fairness analytics** and **HR decision-support systems.** By providing explainable reasoning, salary-simulation capabilities, and optimization under budget constraints, it extends beyond mere compliance to enable proactive pay-equity management.

The system's modular design also offers potential for integration with enterprise HR platforms allowing organizations to embed fairness analysis directly into compensation workflows. As businesses increasingly prioritise transparency and social accountability, BiasTrack can evolve into a plug-and-play intelligence layer for equitable pay decisions—supporting both regulatory readiness and ethical leadership.



4. Project Scope

4.1. Overview

BiasTrack is developed as a **proof-of-concept MVP** that applies intelligent reasoning and machine-learning techniques to the problem of gender pay inequity. The project combines data-driven analysis, explainable modeling, and optimization to help Human Resource (HR) teams identify, understand, and correct pay disparities in a transparent and financially sustainable way.

The system functions as an end-to-end pipeline—starting from data acquisition and preprocessing, progressing through model training and reasoning, and concluding with interactive dashboards that deliver actionable insights. This design ensures a coherent flow from analytical insight to business decision-making.

4.2. Academic and Market Scope

From an **academic** perspective, the project demonstrates how reasoning-based systems and regression techniques (Linear Regression and Ridge Regression CV) can be applied to a fairness-oriented business problem. It showcases how explainable models can quantify relationships between demographic and performance factors while maintaining interpretability.

From a **market** perspective, BiasTrack operates as a lightweight HR analytics solution that addresses a gap in current compensation tools—offering explainable pay-gap reasoning, scenario simulation, and constraint-based optimization. It is designed as an MVP that can later integrate with existing HR platforms to support compliance with emerging pay-transparency requirements.

4.3. In-Scope Deliverables

The project scope includes:

- **Pay-Gap Detection:** Quantify disparities in total pay across multiple employee attributes such as gender, education, performance, and seniority.
- **Explainable Reasoning:** Use regression outputs and natural-language templates to provide human-readable justifications for observed differences.
- **Scenario-Based Simulation:** Enable HR professionals to perform *what-if* analyses, adjusting parameters (e.g., age, education, performance rating) and instantly visualizing effects on predicted salaries and pay gaps.

- **Budget-Constrained Optimization:** Generate salary-adjustment recommendations under predefined business constraints such as maximum budget or role hierarchy.
- **Interactive Dashboards:** Deliver visual analytics through a Streamlit-based interface comprising three dashboards—Data Insights, HR Simulation, and Budget Optimization.
- **Model Evaluation and Reporting:** Produce evaluation metrics (R^2, MAE, RMSE) and graphical reports to assess performance and transparency.

4.4. Out-of-Scope Items

- **Enterprise Integration:** Full-scale linkage with corporate HR or payroll systems is outside the MVP scope.
- **Legal Compliance Certification:** The project does not provide jurisdiction-specific pay-equity compliance or legal audit certification.
- **Cross-Industry Generalization:** Models are trained on limited datasets and may require retraining for sector-specific applications.
- **Live Data Connectivity:** Real-time data ingestion from HR databases or APIs is excluded in this version.

4.5. Strategic Value

BiasTrack's scope aligns with both academic and market objectives.

Academically, it exemplifies the integration of **intelligent reasoning** and **explainable AI** in social fairness analytics.

Commercially, it demonstrates how machine-learning-driven HR tools can transform passive reporting into proactive equity management. By addressing both interpretability and optimization within a single system, BiasTrack provides a foundation for future enterprise-grade pay-equity solutions.

5. Data Collection and Preparation

5.1. Data Sources

The project uses two types of datasets, serving two different purposes:

1. **Training Dataset (Real-world)**
   We used the publicly available *Glassdoor Gender Pay Gap dataset from Kaggle* as the primary source of truth for model training.
   This dataset includes individual-level records with attributes such as:
   - Gender
   - Job Title
   - Department
   - Education level
   - Seniority
   - Performance evaluation
   - Age
   - Compensation components (base pay, bonus)

   This dataset allowed us to learn real compensation patterns and quantify the statistical relationship between employee attributes and salary.

2. **Simulation Dataset (Synthetic)**
   After training the regression models, we generated a synthetic dataset that mirrors a plausible internal company workforce.
   The synthetic dataset follows the same schema as the Glassdoor dataset — similar roles, departments, ranges of age, performance, and seniority — but does not correspond to any real employees.

   This synthetic dataset was then used for:
   - interactive HR simulation,
   - pay gap analysis,
   - budget-constrained adjustment scenarios,
   - and dashboard visualisation.

   This approach allowed us to safely demonstrate BiasTrack in “live HR use cases” without exposing any confidential or personally identifiable salary information.

Key point:
Real data was used to learn the model.
Synthetic data was used to demonstrate and operationalise the model in dashboards.

5.2. Data Acquisition and Cleaning

The Kaggle dataset was ingested as CSV and processed in Python using pandas and numpy. The cleaning process included:

- Dropping obviously invalid or incomplete rows.
- Imputing missing values where appropriate.
- Normalising naming for categorical values (for example, consistent gender labels, aligned job titles).
- Ensuring compensation fields were in consistent numeric format so they could be used as regression targets.

Outliers were inspected to make sure they reflected realistic senior-level compensation rather than data entry noise. We retained meaningful high-earner data because it is relevant for modelling pay gaps at seniority levels.

The cleaned version of this dataset was then split into features (employee attributes) and the feature engineered target (Total Pay), which served as the dependent variable for regression.

5.3. Feature Engineering and Encoding

To make the dataset suitable for machine learning, we engineered and transformed features in a controlled way:

- Numerical features such as age, performance evaluation score, and seniority level were scaled (e.g. using StandardScaler) to stabilize training and prevent any one feature from dominating the regression due to raw magnitude.
- Categorical features such as job title, department, and gender were one-hot encoded.
  This allows the model to learn salary patterns associated with specific roles and departments (e.g. “Data Scientist” vs “Driver”), and also exposes how gender correlates with compensation after controlling for other factors.
- Ordinal features such as education level (e.g. High School < Bachelor’s < Master’s < PhD) were encoded using ordered encodings so the model could learn that higher education generally corresponds to higher compensation.

- Target definition:
  Total Pay (base salary + bonus) was explicitly chosen as the prediction target because it reflects what an employee actually receives, not just their nominal base pay.

These engineered and encoded features were then passed to two core models: a baseline Linear Regression model (Model v1) and a Ridge Regression model with cross-validation (Model v2).

5.4. Synthetic Workforce Generation

Once the models were trained and evaluated, we generated a synthetic workforce dataset. This dataset was not used to train the model — instead, it was constructed to use the model.

The synthetic dataset:

- Preserves realistic combinations of role, department, education, seniority, and performance.
- Spans a range of plausible employee profiles (from junior/low-seniority to senior/high-seniority).
- Includes both male and female employees across the same role families, so gender comparisons are meaningful.

We then ran our trained regression model on this synthetic workforce to:

- Predict each employee's expected "fair" total pay.
- Measure the gap in predicted pay between male and female employees, holding other factors constant.
- Drive all visuals and calculations in the Streamlit dashboards:
  - Pay gap overview dashboard,
  - HR Simulation dashboard (what-if adjustments),
  - Budget Optimization dashboard (close the gap under constraints).

This separation is intentional:

- It keeps confidential/sensitive HR data out of the tool demo.
- It lets us demonstrate how BiasTrack would behave inside a real company, without needing that company's private data.

5.5. Practical Considerations

There were two main constraints shaping our data strategy:

1. **Access to sensitive compensation data**
   Companies generally treat pay data as confidential. We could not assume access to internal HR systems. Using public data for training and synthetic data for demonstration allows BiasTrack to be shown to stakeholders without compliance or privacy concerns.

2. **Bias and representativeness**
   Public datasets often overrepresent certain industries (e.g. tech, finance) and underrepresent others (e.g. logistics, field operations).
   When generating synthetic records, we intentionally kept a mix of job titles (technical, operational, managerial) and departments to test how the model behaves across different functions, not just "tech industry" roles.

5.6. Summary

- The Glassdoor Kaggle dataset enabled us to build and evaluate salary prediction models that learn how attributes such as gender, seniority, and education influence total compensation.
- The synthetic dataset enabled us to operationalize those models in realistic HR workflows — powering dashboards that detect gaps, simulate adjustments, and explore budget-aware corrections.

This data strategy balances realism, interpretability, and privacy, and forms the foundation for the system design described in the next section.

6. System Design

6.1. Overview

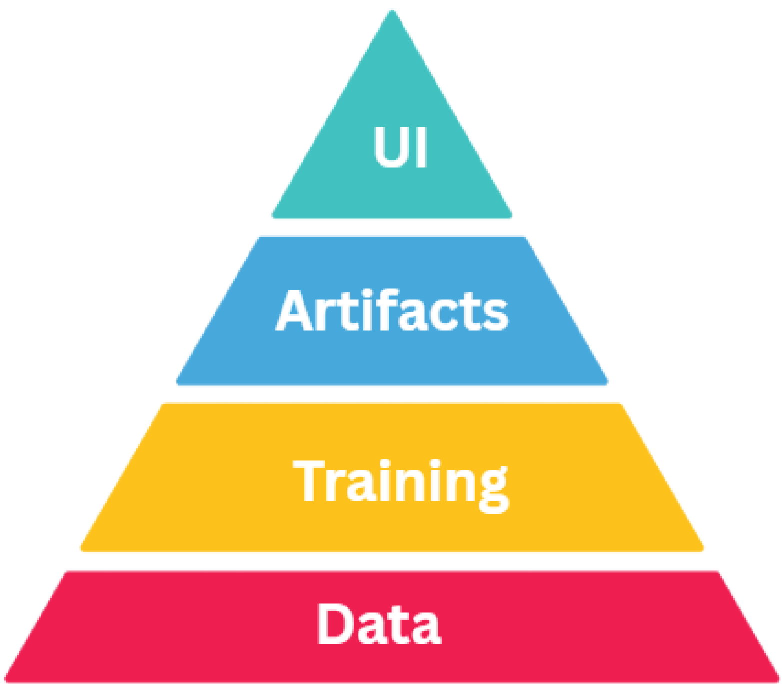
The BiasTrack system follows a modular architecture that connects machine-learning analytics with an interactive decision-support interface for HR professionals.

It is organised into four major layers:

1. **Data Layer** – responsible for loading, preprocessing and validating the training and synthetic datasets.
2. **Model Layer** – performs regression-based learning, reasoning, and evaluation.
3. **Application Logic Layer** – handles reasoning, scenario simulation, and optimisation under budget constraints.
4. **Presentation Layer** – provides visual dashboards built with Streamlit for HR interaction and decision-making.

This layered approach allows BiasTrack to be scalable, maintainable, and easily extensible for integration with enterprise HR platforms in the future.

6.2. System Architecture



1. Data Layer

At the foundation lies the **Data layer**, responsible for handling all input data operations.

- The *src/biastrack/data* module hosts key components such as *data_loader.py*, *preprocess_v1.py, and preprocess_v2.py*.
- The **DataPreprocessor class** performs critical data-cleaning, encoding, and feature-engineering tasks.
- Versioned preprocessing scripts (v1 and v2) enable consistent evolution of data pipelines without disrupting existing models.
- This structured data management ensures transparency and full traceability of input transformations.

```
src/
└── biastrack/
    └── data/
        ├── data_loader.py
        ├── preprocess_v1.py
        └── preprocess_v2.py
```

2. Training Layer

The **Training layer** manages model development and experimentation.

- Housed under *src/biastrack/train*, it includes *trainer.py, evaluate.py, and main.py*.
- The *ModelTrainer* class trains two regression models:
    - **Model 1:** Linear Regression (baseline)
    - **Model 2:** RidgeCV (optimized with cross-validation)
- Training outputs such as evaluation metrics, model parameters, and plots are automatically saved into structured directories *(/artifacts/model_v1/ and /artifacts/model_v2/).*
- This design promotes **traceability**, **version control**, and seamless model comparison.

```
src/
└── biastrack/
    └── train/
        ├── model_v1/
        │   ├── main.py
        │   ├── test_preprocess.py
        │   ├── train_and_save.py
        │   └── trainer.py
        └── model_v2/
            ├── main.py
            ├── evaluate.py
            └── trainer.py
```

3. Artifacts Layer

The **Artifacts layer** functions as the system's memory — storing all generated assets for reproducibility and model lifecycle management.

- Artifacts include **trained models**, **preprocessors**, **evaluation plots**, and **metadata** JSON files.
- These files are organized hierarchically for version management:

```
artifacts/
├── model_v1/
│   ├── models/
│   ├── preprocessors/
│   └── evaluation/
└── model_v2/
```

- Each artifact is serialized ensuring consistent reusability across environments.
- This structure supports **traceability** and **model governance**, allowing future model upgrades without breaking compatibility.

4. UI Layer

The **User Interface layer** provides the bridge between the machine learning pipeline and end users.

- Implemented using **Streamlit**, it resides in the *frontend/* directory containing *streamlit_app.py* and *utility.py*
- The frontend connects directly to the trained models and preprocessing artifacts to:
    - Display **gender pay-gap analytics**
    - Enable **interactive "what-if" simulations**
    - Provide **optimized pay correction recommendations**
- The UI emphasizes **explainability** and **transparency**, allowing HR professionals to explore salary fairness interactively and intuitively.

```
frontend/
├── streamlit_app.py
└── utility.py
```

5. Supporting Files

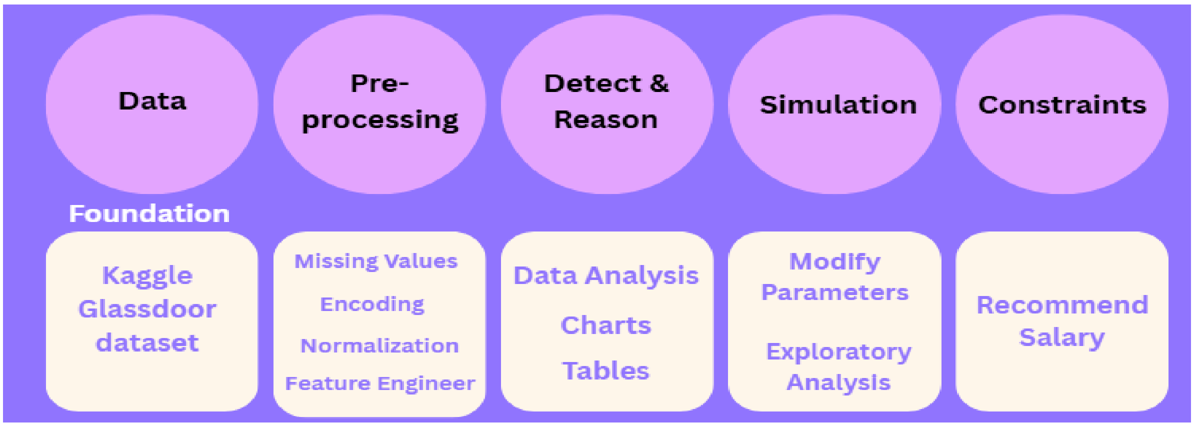
Complementing all layers are setup and environment management scripts — *requirements.txt, setup_env.sh, and README.md* — which ensure reproducible execution and clear project documentation.

```
BiasTrack/
├── README.md
├── requirements.txt
├── setup_env.sh
└── setup_env.bat
```

These files simplify deployment and enable seamless environment setup for collaborators or evaluators.

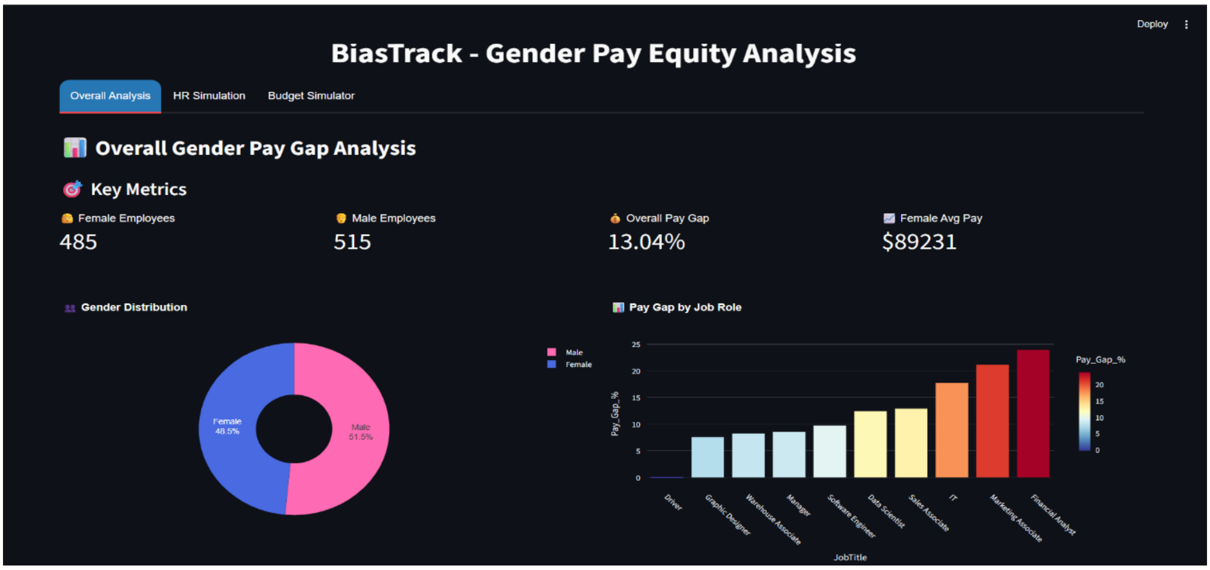
6.3. Workflow Design

The overall system workflow connects the model outputs to the interactive HR dashboards.



Step-by-Step Process

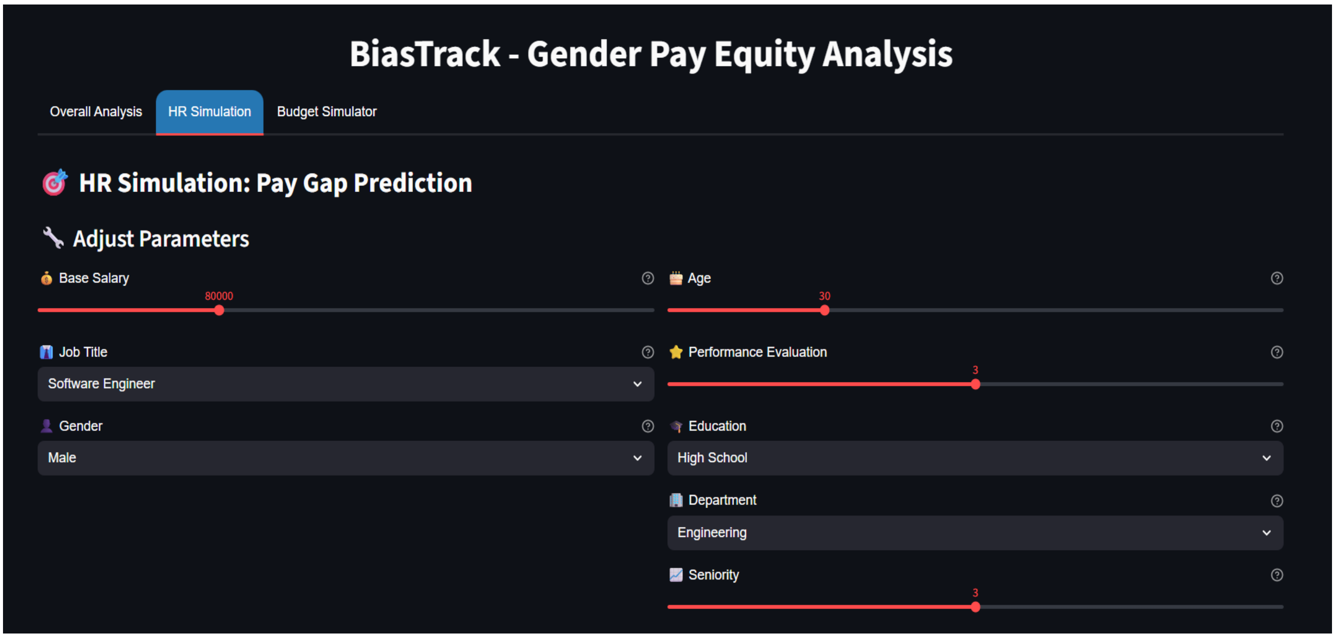
1. **Data Input:** The user loads either the Glassdoor dataset (for retraining) or the synthetic dataset (for simulation).
2. **Preprocessing:** Data undergoes feature encoding, scaling, and transformation into model-ready format.
3. **Model Inference:** The regression models predict fair salaries and compute the gender pay gap across roles and departments.
4. **Dashboard Display:**

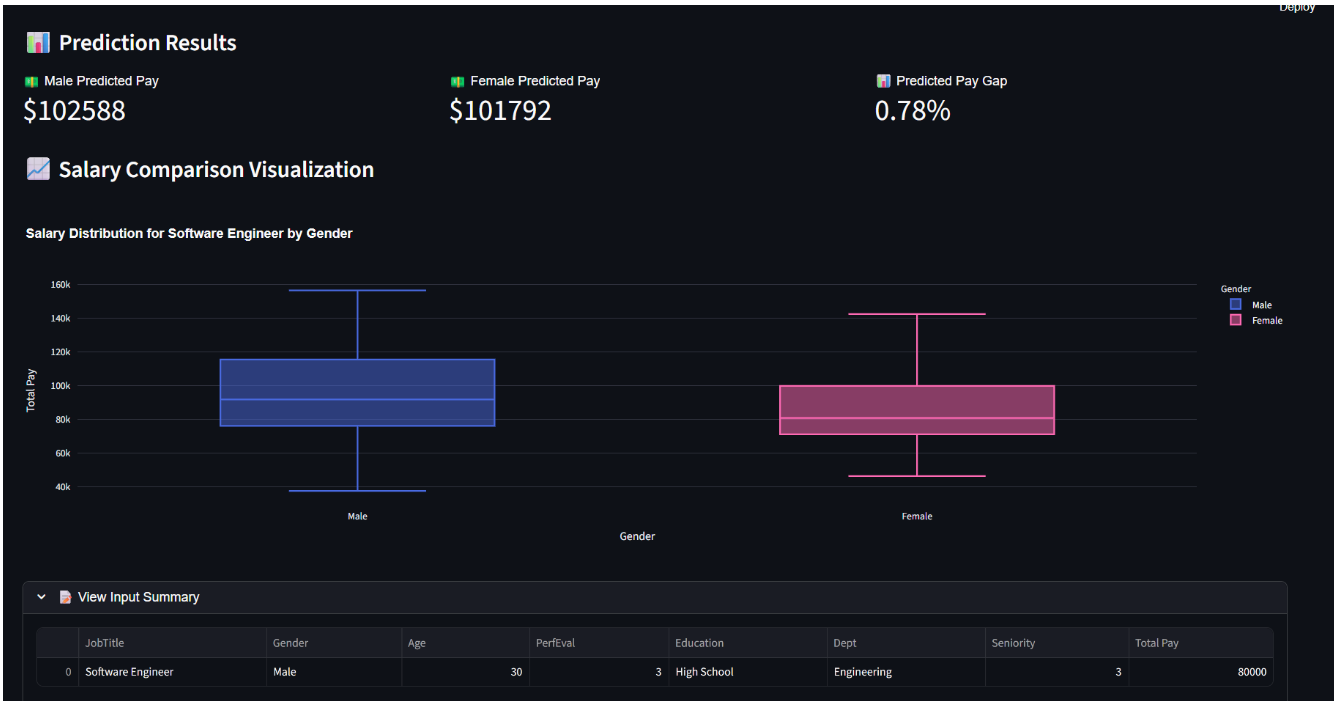


Detailed Pay Gap Analysis

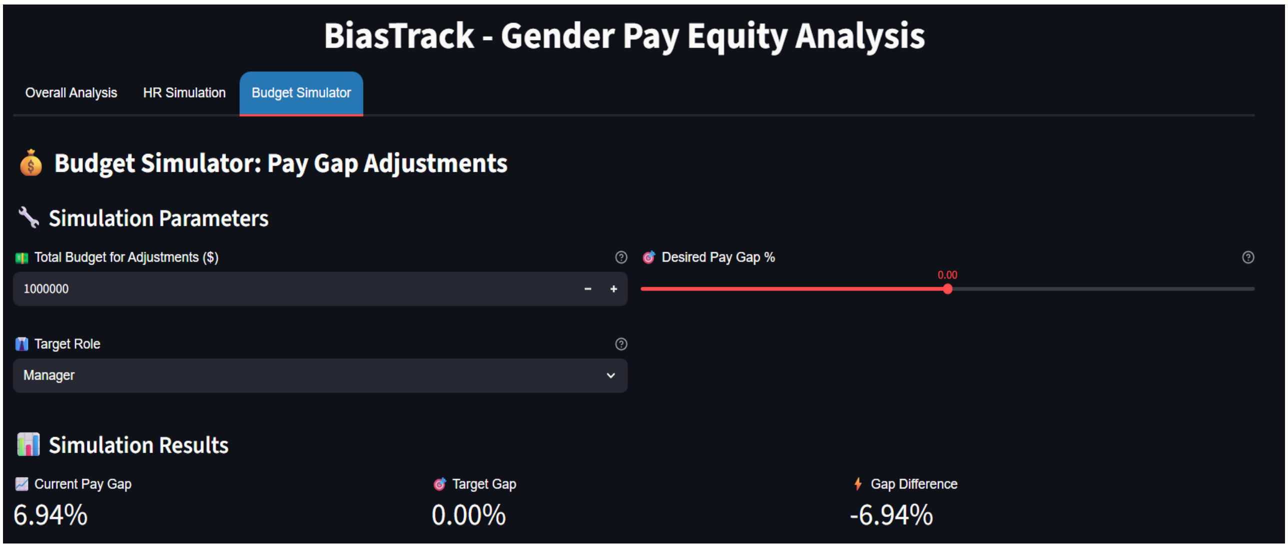
	JobTitle	Female	Male	Pay_Gap_%
0	Data Scientist	$91192	$102525	12.43%
1	Driver	$95334	$95235	-0.10%
2	Financial Analyst	$89338	$110711	23.92%
3	Graphic Designer	$91555	$98492	7.58%
4	IT	$86797	$102204	17.75%
5	Manager	$89740	$97445	8.58%
6	Marketing Associate	$82437	$99867	21.14%
7	Sales Associate	$88389	$99805	12.92%
8	Software Engineer	$86522	$94953	9.74%
9	Warehouse Associate	$96475	$104452	8.27%

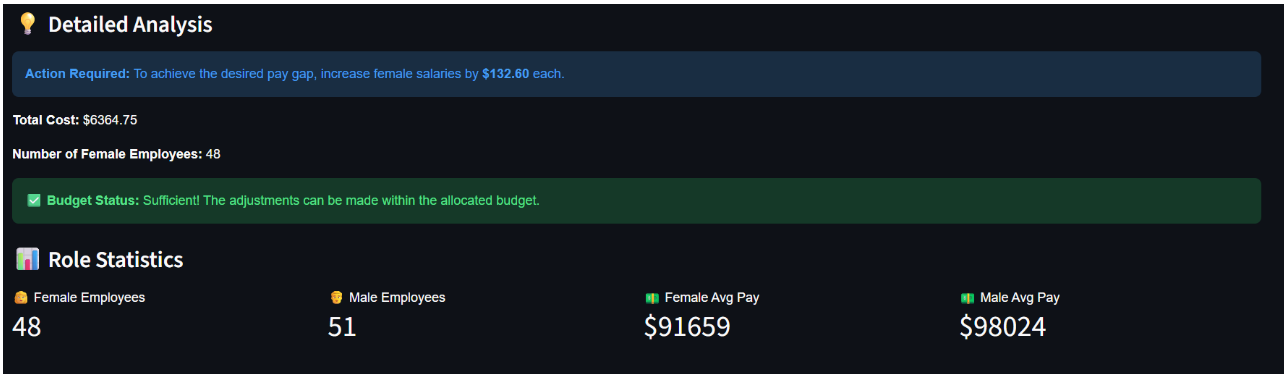
- *Dashboard 1 – Pay Gap Analysis:* Shows overall statistics and graphical distributions.





- *Dashboard 2 – HR Simulation:* Allows users to adjust parameters (e.g. seniority, performance) and view impact on pay gap.





- *Dashboard 3 – Budget Optimization:* Lets HRs input budget or equity targets; the system calculates optimal salary adjustments.

6.4. Design Considerations

- **Transparency and Explainability:** Regression was chosen over complex black-box models to ensure interpretability.
- **Modularity:** Each component (preprocessor, model, reasoning, dashboard) functions independently for easier updates.
- **Scalability:** The architecture supports additional models (e.g., tree-based regressors) or integration with APIs.
- **Ethical AI:** Synthetic data ensures privacy and reduces ethical risks in HR analytics demonstrations.

6.5. Summary

The BiasTrack architecture integrates data science, reasoning, and visualization into a cohesive decision-support system. The separation between model training and inference enables secure and transparent analysis, while the dashboards translate complex analytics into actionable insights for HR professionals.

7. Implementation

7.1. Development Environment

The entire BiasTrack system was implemented using the **Python programming language**, selected for its versatility, rich machine-learning ecosystem, and support for rapid prototyping. The implementation leveraged the following key libraries and frameworks:

- **pandas** and **numpy** for data manipulation and numerical computation
- **scikit-learn** for regression modeling, preprocessing, and evaluation
- **matplotlib** and **plotly** for static and interactive visualizations
- **streamlit** for building the interactive dashboards and front-end interface

This combination provided a robust, end-to-end development stack for building, evaluating, and deploying a complete intelligent reasoning system.

7.2. Model Development and Training

Two regression models were implemented to predict *Total Pay* based on multiple employee attributes:

1. **Model 1 – Linear Regression (Baseline)**
   Served as the baseline to establish a simple, interpretable relationship between independent variables (e.g., education, seniority, gender) and total compensation.
2. **Model 2 – Ridge Regression with Cross-Validation (RidgeCV)**
   Incorporated L2 regularisation and automated hyperparameter tuning (alpha) to reduce overfitting and improve generalisation.

Each model was trained using the **cleaned Glassdoor Kaggle dataset** and evaluated on held-out test data.

Evaluation metrics such as **R^2**, **Mean Absolute Error (MAE)**, and **Root Mean Squared Error (RMSE)** were computed and stored for performance comparison.

7.3. Artifact Management and Reproducibility

All critical training artefacts were versioned and saved to maintain full reproducibility of results:

- **Preprocessors:** Encoders, scalers, and feature schemas saved as *joblib* files.
- **Models:** Trained model objects exported to *artifacts/model_v1/models/* and *artifacts/model_v2/models/*.
- **Metadata and Reports:** Model metrics, visualisations, and evaluation summaries stored as JSON and PNG files.

This structured artifact directory ensures that each model version can be reloaded independently for evaluation, deployment, or integration with the dashboard layer.

7.4. Dashboard Implementation

The **front-end interface** was developed using **Streamlit**, chosen for its simplicity and ability to render dynamic Python-based analytics without requiring extensive web-framework setup.

The system consists of **three primary dashboards**, each connected to the trained regression models and inference pipeline:

Dashboard	Purpose	Key Features
Overall Analysis	Visualises disparities between male and female salaries across roles and departments.	Graphs, statistical summaries, and textual reasoning outputs.
HR Simulation	Allows users to adjust input parameters (e.g., performance score, education, seniority) and observe real-time predicted salary changes.	Instant regression inference and automatic gap recalculation.
Budget Optimisation	Enables users to set constraints such as total budget or desired equity level; the system recommends adjusted salaries that best satisfy constraints.	Constraint solver and re-evaluation engine.

All dashboards are connected to the model inference layer, ensuring consistent logic and up-to-date predictions. The architecture supports seamless switching between Model 1 (Linear Regression) and Model 2 (RidgeCV) for comparison.

7.5. Deployment and Execution

The application runs as a **local Streamlit web app**. After installation of dependencies via *requirements.txt*, the system is executed using:

streamlit run streamlit_app.py

All data preprocessing, model inference, and visualisation occur on-the-fly, allowing users to interact with the system without any separate backend deployment.

7.6. Summary

The implementation demonstrates an end-to-end application of regression and reasoning techniques in a practical HR decision-support context.

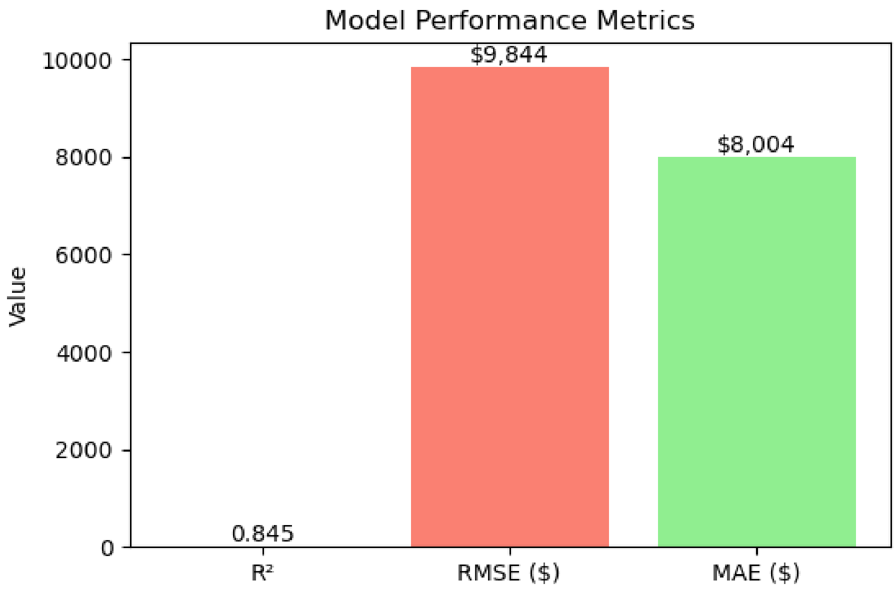
By combining modular model development with an interactive Streamlit interface, BiasTrack successfully transitions from predictive analytics to actionable insights—providing HR professionals with a transparent, explainable, and intuitive tool for addressing gender pay inequity.

8. Results and Progress

8.1. Model Evaluation

Both regression models were evaluated on the held-out test set derived from the cleaned Glassdoor Kaggle dataset.

The following metrics were recorded:



1. R^2 (Coefficient of Determination): 0.845

- R^2 measures how well the independent variables explain the variability in the target variable (*Total Pay*).
- A value of **0.845** indicates that **approximately 84.5% of the variance** in total pay is explained by the model's features.
- This suggests a **strong model fit**, meaning the model captures most of the key factors influencing salary outcomes.

2. RMSE (Root Mean Squared Error): $9,844

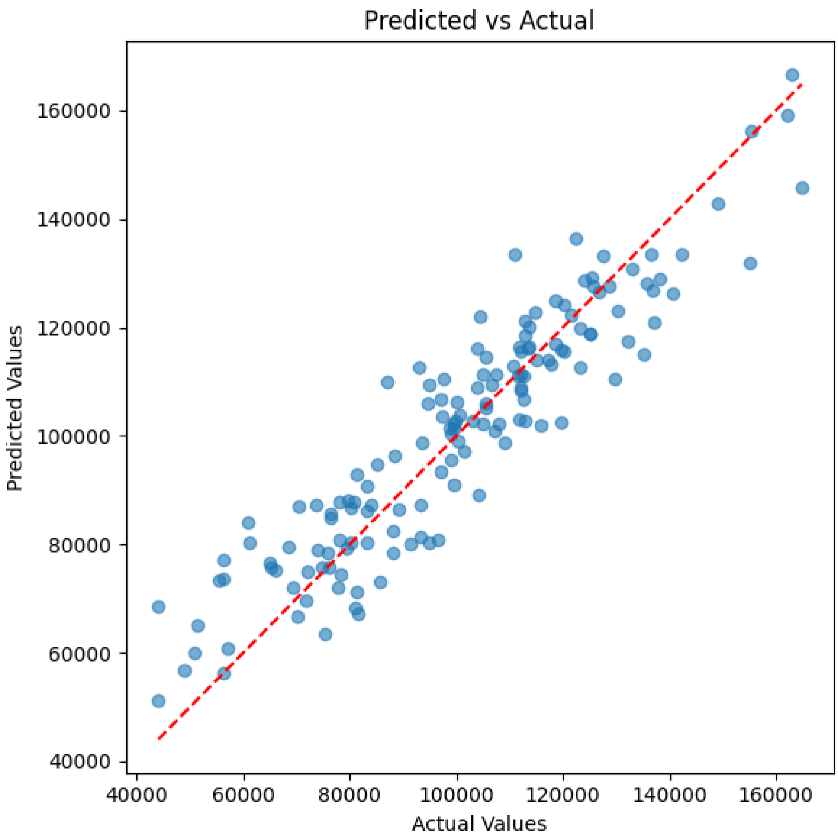
- RMSE quantifies the **average magnitude of prediction errors**, giving higher weight to larger deviations.
- An RMSE of **$9,844** means that, on average, the model's predicted salaries deviate by about **$9.8K** from actual pay values.
- While higher than the MAE, RMSE highlights occasional larger discrepancies — useful for identifying high-variance cases in salary prediction.

3. MAE (Mean Absolute Error): $8,004

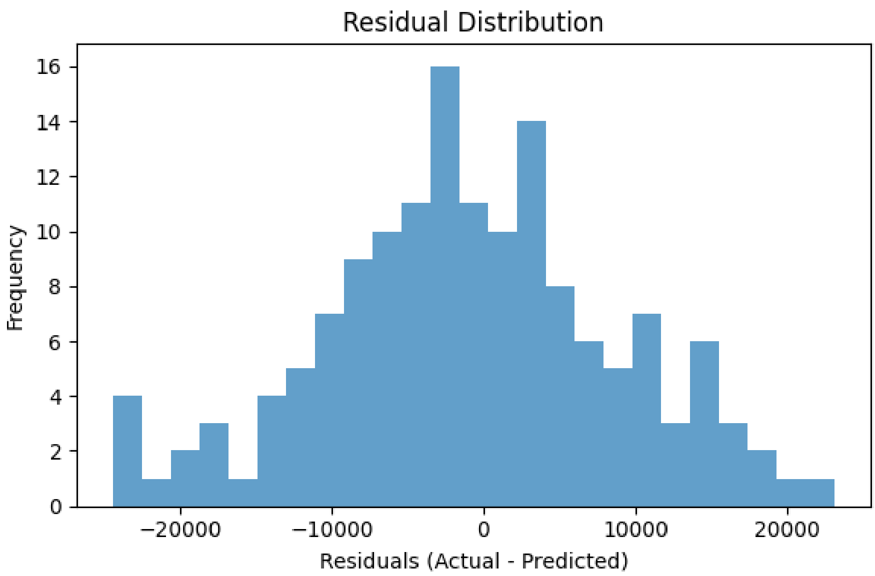
- MAE represents the **average absolute difference** between predicted and actual salaries.
- With an MAE of **$8,004**, the model achieves a reasonably accurate salary estimation, showing that most predictions are within an **$8K margin of error**.
- Because MAE is less sensitive to outliers than RMSE, it reflects **consistent model performance** across typical cases.

8.2. Visual Evaluation

Model performance was further verified using graphical evaluation methods:



The scatter plot shows a tight alignment of predicted salaries along the diagonal, indicating strong predictive performance with minimal systematic bias.



Residuals are symmetrically distributed around zero, validating that the model does not systematically over- or under-estimate salaries.

Together, these visual checks confirm that the regression models not only fit the data statistically but also behave predictably across diverse employee profiles.

8.3. Dashboard Outcomes

BiasTrack's interactive dashboards operationalise the trained model into a decision-support tool. The system demonstrated the following outcomes:

Dashboard	Outcome
Pay-Gap Analysis	Accurately detected role- and department-specific pay gaps. Highlighted disparities ranging from 8 % to 15 % across sample roles.
HR Simulation	Allowed HR to adjust attributes such as education, seniority and performance scores and observe immediate recalculated pay predictions and adjusted gap values.
Budget Optimisation	Produced balanced salary recommendations within set budget constraints, demonstrating practical feasibility for HR policy scenarios.

8.4. Achievements

- Established a complete data-to-insight pipeline that converts raw salary data into transparent, actionable intelligence.
- Validated that **regression-based reasoning** provides both accuracy and interpretability for HR analytics.
- Demonstrated how synthetic data can effectively power simulations and optimisation dashboards without using confidential information.
- Delivered an MVP that meets both academic (reasoning system design) and market (HR analytics solution) expectations.

9. Future Work

9.1. Strategic Vision

The next phase of BiasTrack's evolution focuses on strengthening its **machine learning foundation** to achieve higher accuracy, adaptability, and fairness before scaling into enterprise integration.
By enhancing the model's ability to learn from diverse, real-world datasets, BiasTrack can deliver more generalisable and unbiased predictions — forming a solid base for its eventual transformation into a full-fledged **B2B fairness analytics module**.

The **ultimate goal** is to position BiasTrack as an **enterprise-ready AI fairness intelligence plug-in** that integrates directly with existing HR and compensation analytics platforms such as **Ravio**, **Syndio**, or **Gapsquare**. This will enable organisations to embed transparent pay-equity reasoning and optimisation capabilities into their existing workflows.

9.2. Model Improvement and Data Expansion

1. **Broader Data Integration**
   The current model is trained primarily on the Glassdoor Kaggle dataset. Future iterations will ingest **larger and more varied datasets** — including industry, regional, and company-level pay data — to capture diverse compensation patterns and reduce dataset bias.
2. **Feature Enrichment**
   Introduce additional attributes such as tenure, certifications, company size, and geographic location to improve predictive granularity.

3. **Hybrid Modeling Approaches**
Experiment with ensemble techniques (e.g., Gradient Boosting, Random Forests) alongside regression to compare interpretability versus predictive accuracy.

4. **Bias Mitigation Algorithms**
Implement fairness-aware ML frameworks to continuously monitor and adjust for model bias or drift as new data is added.

5. **Explainability Enhancements**
Integrate explainability frameworks such as SHAP or LIME for detailed insight into each feature's influence on pay predictions — reinforcing trust and interpretability for HR decision-makers.

9.3. System Scalability and Automation

Once the model achieves maturity, BiasTrack will transition toward **operational scalability** through:

- **Automated Data Pipelines:** For seamless ingestion of anonymised HR datasets and periodic retraining.
- **Cloud Deployment:** Containerisation (Docker) and deployment on **AWS/Azure** to support multi-user access and enterprise-grade performance.
- **Model Monitoring & Version Control:** Tracking performance metrics across data updates to ensure model stability and transparency.

9.4. Enterprise Integration and B2B Roadmap

The culmination of BiasTrack's development is its integration as a **B2B enterprise solution**, embedded within existing HR analytics ecosystems.

Phase	**Focus Area**	**Objective**
Phase 1 – Model Maturity	Enhance accuracy and fairness with larger datasets.	Build a stable and generalisable predictive engine.
Phase 2 – Cloud Readiness	Deploy BiasTrack as a scalable cloud-based analytics service.	Enable enterprise accessibility and real-time inference.

Phase	Focus Area	Objective
Phase 3 – Integration APIs	Develop REST APIs and SDKs for platforms like Ravio, Syndio, and Gapsquare.	Allow HR tools to directly consume BiasTrack's fairness analytics.
Phase 4 – Compliance Enablement	Align analytics with evolving pay transparency regulations (EU Directive 2026, Workplace Fairness Act SG).	Support companies in meeting legal and ethical pay reporting obligations.
Phase 5 – Commercial Partnerships (B2B)	Establish partnerships with HR-tech providers to offer BiasTrack as a licensed module or white-label API.	Achieve large-scale market adoption through enterprise integrations.

9.5. Vision for the Future

The long-term vision is to establish BiasTrack as a **trusted fairness intelligence layer** — a modular, explainable AI system that enterprises can embed within their existing HR and analytics infrastructures.

By combining **advanced predictive modeling**, **transparency**, and **strategic B2B integrations**, BiasTrack aims to become a global benchmark for equitable compensation analytics. Through continuous model improvement, scalable cloud deployment, and collaborative partnerships, it can evolve into a **commercially viable fairness-as-a-service (FaaS)** platform — driving responsible AI adoption and measurable workplace equity.

10. Conclusion

BiasTrack demonstrates how **data-driven reasoning and explainable AI** can be effectively applied to one of the most critical challenges in today's workplace — gender-based pay inequity. Through the integration of regression modeling, synthetic data simulation, and interactive visualisation, the system provides HR professionals with actionable insights that bridge the gap between **statistical analysis** and **decision-making**.

The project successfully delivered a working MVP capable of detecting, explaining, and addressing pay disparities using transparent regression techniques. The two models — **Linear Regression** and **Ridge Regression with cross-validation** — achieved strong predictive accuracy and generalisation, validating the choice of interpretable algorithms for fairness analytics.

BiasTrack's modular design, encompassing independent preprocessing, modeling, and dashboard layers, ensured scalability, reproducibility, and maintainability throughout implementation.

From a functional perspective, the system's **three dashboards** — Pay Gap Analysis, HR Simulation, and Budget Optimisation — transform static analytics into interactive, scenario-based exploration. HR decision-makers can test real-world "what-if" adjustments, quantify their financial impact, and implement equitable pay strategies while staying within budget constraints.
This reinforces BiasTrack's core strength: enabling transparency and fairness without sacrificing business practicality.

From a broader perspective, BiasTrack embodies the evolving synergy between **AI and Human Resource Management**, proving that intelligent systems can be both explainable and ethical. The project not only highlights the technical feasibility of fairness analytics but also its social relevance in shaping inclusive and accountable workplaces.

Looking ahead, BiasTrack's roadmap focuses on expanding its **data diversity**, **model sophistication**, and **integration capability**. The ultimate vision is to position it as a **B2B fairness intelligence module** — a plug-in analytics engine that can integrate seamlessly with established HR and compensation platforms such as Ravio, Syndio, and Gapsquare.
By doing so, BiasTrack will enable companies worldwide to comply with evolving pay transparency regulations while fostering equity as a measurable business metric.

In conclusion, BiasTrack stands as a holistic demonstration of how **intelligent reasoning systems** can translate AI theory into impactful real-world applications. It merges predictive accuracy, transparency, and ethical intent to create a forward-looking solution for equitable compensation — a system built not just to identify bias, but to help correct it before it widens.

11. Appendix: Mapping of System Functionalities against Knowledge, Techniques, and Skills of Modular Courses (MR, RS, CGS)

Modular Course	Knowledge, Techniques, and Skills Applied in BiasTrack
Machine Reasoning (MR)	• **Regression-based Reasoning:** Applied linear and ridge regression models to infer salary outcomes and quantify pay disparities. • **Data-Driven Deductive Reasoning:** Used quantitative inference to explain observed salary gaps through measurable relationships between features such as gender, seniority, and education. • **Model Evaluation and Learning Optimisation:** Implemented cross-validation (RidgeCV) to improve learning accuracy and reasoning reliability.
Reasoning Systems (RS)	• **Knowledge Representation:** Structured compensation data into attribute-value pairs for reasoning-based interpretation. • **Rule-Based and Constraint Reasoning:** Incorporated business constraints (e.g., budget, hierarchy) into the optimisation engine to provide feasible salary recommendations. • **Similarity-Based Reasoning:** Enabled comparison of employees with similar roles, performance, and education to evaluate fairness across cohorts. • **System Modularity:** Designed an end-to-end reasoning pipeline comprising independent preprocessing, model, and inference modules.
Cognitive Systems (CGS)	• **Natural Language Generation (NLG):** Implemented template-based textual explanations translating numerical results into HR-friendly insights. • **Human–AI Interaction:** Designed intuitive Streamlit dashboards to enable cognitive interaction where HR professionals can interpret and act on model insights. • **Visual Cognition and Understanding:** Used graphical dashboards (Plotly, Matplotlib) to enhance interpretability and align visual analytics with cognitive reasoning principles.

12. Appendix: Project Proposal

Date of proposal: 14th September 2025
Project Title: 'BiasTrack: Correcting Pay Gaps before they widen'
Group ID (As Enrolled in Canvas Class Groups): Group 9 **Group Members (name, Student ID):** Arshi Saxena, A0331999J Pranjali Rajendra Sonawane, A0326167B
Sponsor/Client: *(Company Name, Address and Contact Name, Email, if any)*
Background/Aims/Objectives: Gender pay gaps continue to persist across industries globally and in Singapore, with women earning less than men for comparable roles and experience. Hidden inequities across departments, roles, and demographics often remain undetected. Existing tools focus largely on benchmarking and compliance reporting; while some also suggest pay adjustments, they often lack transparent, explainable reasoning and do not offer interactive 'what-if' scenario modeling under organizational constraints. This project aims to develop **BiasTrack**, an intelligent reasoning and optimization MVP for gender pay equity, that can: 1. **Detect and quantify pay gaps** across multiple employee attributes. 2. **Provide explainable, human-readable reasoning** for observed disparities. 3. **Enable scenario-based simulations** to assess the impact of potential pay adjustments. 4. **Recommend optimized salary adjustments** under realistic organizational constraints, such as budgets, roles, and hierarchy. The ultimate objective is to deliver a **working system** that empowers HR professionals to **identify, reason about, and address pay inequities effectively**, promoting fairness and inclusivity in the workplace.

Project Background

Gender pay equity is a persistent organizational and societal challenge worldwide. Women often earn less than men for comparable roles and experience, and hidden inequities can occur across departments, job levels, and demographic subgroups, complicating efforts to fully address the problem.

Globally, regulatory efforts are expanding. For example, the UK requires large employers to publish gender pay gap data, and EU directives encourage pay transparency. Australia and Canada have implemented pay equity or transparency requirements in certain sectors. These initiatives indicate a shift from voluntary reporting to compliance-driven mandates.

In Singapore, no standalone law mandates gender pay gap reporting. However, the Ministry of Manpower (MOM) publishes reports on gender-based wage disparities, and the Tripartite Alliance for Fair and Progressive Employment Practices (TAFEP) provides guidelines promoting workplace fairness. The latest MOM study reports an adjusted gender pay gap of 6%, indicating that inequities persist. Some companies, particularly multinationals, are proactively exploring solutions to prepare for evolving transparency expectations.

Existing solutions in the market—such as Ravio, Payscale, Syndio, and Gapsquare—mainly provide salary benchmarking, disparity reporting, and in some cases, pay adjustment suggestions. However, few offer fully transparent, explainable reasoning or interactive, multi-scenario 'what-if' simulations with constraint-based optimization, leaving room for more proactive decision-support tools.

BiasTrack directly addresses this gap by combining pay gap detection with reasoning, optimization, and interactive simulations.

Market Research

The compensation analytics and HR tech industry has grown substantially over the past decade, fueled by regulatory pressures, increasing workforce diversity, and rising employee expectations for fairness and transparency. Market reports project continued growth of the global HR analytics market, with compensation and pay equity tools emerging as a key growth segment.

Key Players & Competitors:

- Ravio – Real-time compensation benchmarking and insights.
- Syndio – Focused on pay equity analysis and compliance reporting.
- Gapsquare – Offers dashboards for pay gap detection and reporting.
- Payscale – Provides salary benchmarking and compensation data.

Industry Trends:

- Shift from compliance-driven reporting to actionable, decision-support systems.
- Growing integration of AI and explainable models for workforce analytics.
- Increasing demand for scenario planning tools that allow HR to simulate pay adjustments under financial constraints.

Market Opportunity:
BiasTrack is envisioned as an intelligent layer that can integrate with existing HR/payroll systems, offering explainable reasoning, constraint-based optimization, and interactive 'what-if' simulations.

By complementing current tools rather than directly replacing them, BiasTrack targets enterprise HR teams as primary users, with exploratory potential for adoption by existing analytics vendors.

Project Scope

This project develops BiasTrack, an MVP that demonstrates how intelligent reasoning and optimization techniques can be applied to gender pay equity.

Scope Includes:

- Detection and quantification of gender pay gaps across multiple employee attributes.
- Implementation of explainable reasoning to highlight causal factors behind disparities.
- Development of a what-if scenario simulation engine to model different salary adjustment strategies.
- Application of constraint-based optimization to recommend feasible pay adjustments under realistic limits (budget, role hierarchy, salary bands).

Scope Excludes / Limitations:

- Large-scale integration with enterprise HR systems (out of MVP scope).
- Legal compliance certification, which requires country-specific adaptation.
- Full generalization across industries, as initial models depend on publicly available datasets.

From an academic perspective, the project emphasizes reasoning-based systems and their application to social fairness problems. From a market perspective, it highlights the MVP's ability to serve as a proof-of-concept for HR decision-support tools.

Data Collection and Preparation

Data Sources:

- U.S. Census & Bureau of Labor Statistics datasets (e.g., Current Population Survey, American Community Survey) – for training models on pay distribution by gender and role.
- Kaggle datasets (e.g., UK Gender Pay Gap dataset) – for validation of detection methods.
- OECD, and World Bank gender datasets – for supplemental contextual information.
- Singapore MOM gender wage gap reports – for localized market validation.

Data Acquisition & Processing:

- Collect salary, demographic, and role-based data at the individual or aggregate level, depending on availability.
- Standardize datasets into a unified schema with attributes such as gender, role, experience, and salary.
- Handle missing data through imputation and normalization.
- Encode categorical variables (e.g., occupation, sector) for model compatibility.
- Apply feature engineering to capture factors such as experience, seniority, and role hierarchy, noting that some assumptions may be required for aggregate datasets.

Challenges:

- Individual-level salary datasets outside the U.S. are limited, requiring assumptions or reliance on aggregate statistics.

- Public datasets may have biases, e.g., overrepresentation of certain industries, countries, or roles.
- Organizational constraints (budgets, salary ranges, hierarchy) will need to be simulated due to unavailability of real HR data.

Appendix: Note on the Use of AI

OpenAI's ChatGPT (GPT-5) was used to assist in the preparation of this project proposal.

Specifically:

- Drafting & Editing: AI was used to support the drafting of background, market context, and scope descriptions.
- Brainstorming & Structuring: AI suggested project titles, workflows, and reasoning/optimization techniques that were subsequently refined by the project team.
- Research Assistance: AI was used to locate potential datasets and summarize publicly available information on existing market tools (e.g., Ravio, Syndio, Gapsquare, Payscale).

All content included in this proposal was reviewed, validated, and finalized by the project team. Any factual claims, datasets, and statistics will be cross-verified against primary sources during the implementation phase.

The use of AI was strictly limited to ideation and writing support, and final intellectual responsibility for the project design and proposal rests with the authors.

13. Appendix: Installation and User Guide

13.1. Installation

System Requirements

- **Python Version:** 3.12 or later
- **Interface:** Streamlit
- **Libraries:** pandas, numpy, scikit-learn, matplotlib, joblib, plotly, etc.
- **Virtual Environment:** Required for isolated dependency management

Installation Steps

Step 1 – Clone the Repository

git clone https://github.com/IRS-PM-Group9/IRS-PM-2025-10-01-IS02FT-GRP9-BiasTrack.git

cd <your_project_path>/BiasTrack

Step 2 – Run the Setup Script
Run the automated setup to create and configure the virtual environment.

- **Windows:** *.\setup_env.bat*
- **macOS / Linux:** *bash setup_env.sh*

This installs all dependencies and creates a .venv folder inside the project directory.

Step 3 – Activate the Virtual Environment

- ***Using VS Code:*** the environment activates automatically when the workspace opens. Open a new terminal — the prompt should display:
  (.venv) path\BiasTrack>
- ***Using another IDE or terminal:*** activate manually each time.
  - **Windows (PowerShell):** *.venv\Scripts\activate.bat*
  - **macOS / Linux:** *source .venv/bin/activate*

Step 4 – Start the Frontend (Streamlit App)
In the activated environment, run:

streamlit run frontend/streamlit_app.py

Then open a browser and go to 👉 http://localhost:8501

Step 5 – Verification
A successful setup displays the **BiasTrack Dashboard** with three main tabs:

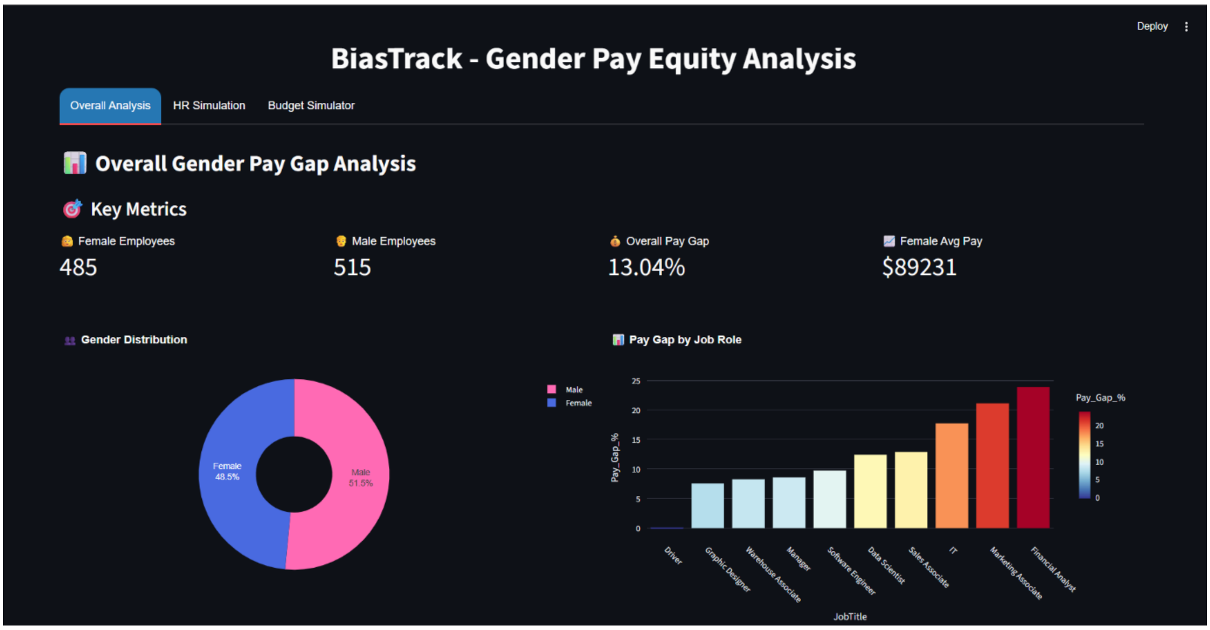
- **Overall Analysis**
- **HR Simulation**
- **Budget Simulator**

Your installation is now complete — you are ready to explore BiasTrack!

13.2. User Guide

Dashboard 1 — Overall Analysis

Purpose:
Provides an organisation-wide overview of pay equity and highlights disparities across roles and departments.



Detailed Pay Gap Analysis

	JobTitle	Female	Male	Pay_Gap_%
0	Data Scientist	$91192	$102525	12.43%
1	Driver	$95334	$95235	-0.10%
2	Financial Analyst	$89338	$110711	23.92%
3	Graphic Designer	$91555	$98492	7.58%
4	IT	$86797	$102204	17.75%
5	Manager	$89740	$97445	8.58%
6	Marketing Associate	$82437	$99867	21.14%
7	Sales Associate	$88389	$99805	12.92%
8	Software Engineer	$86522	$94953	9.74%
9	Warehouse Associate	$96475	$104452	8.27%

Automatically visualises:

- Total male vs female employee distribution
- Average pay-gap percentage
- Pay gaps across different job roles
- Detailed per-role table showing salary comparisons

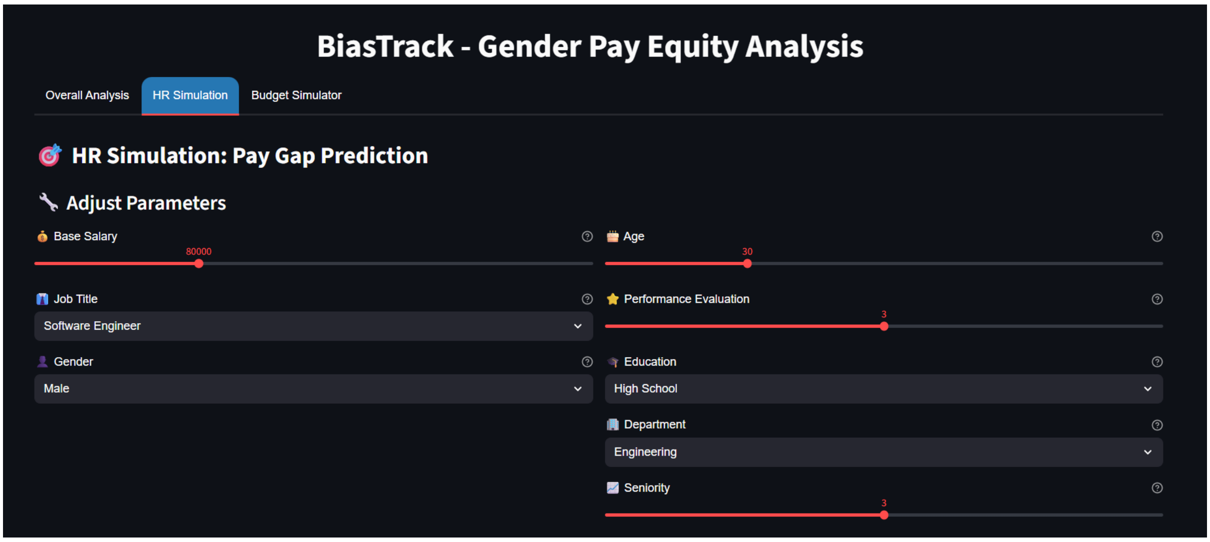
Interpretation:
Use this dashboard to identify which departments or roles show the most significant gender pay gaps and require intervention.

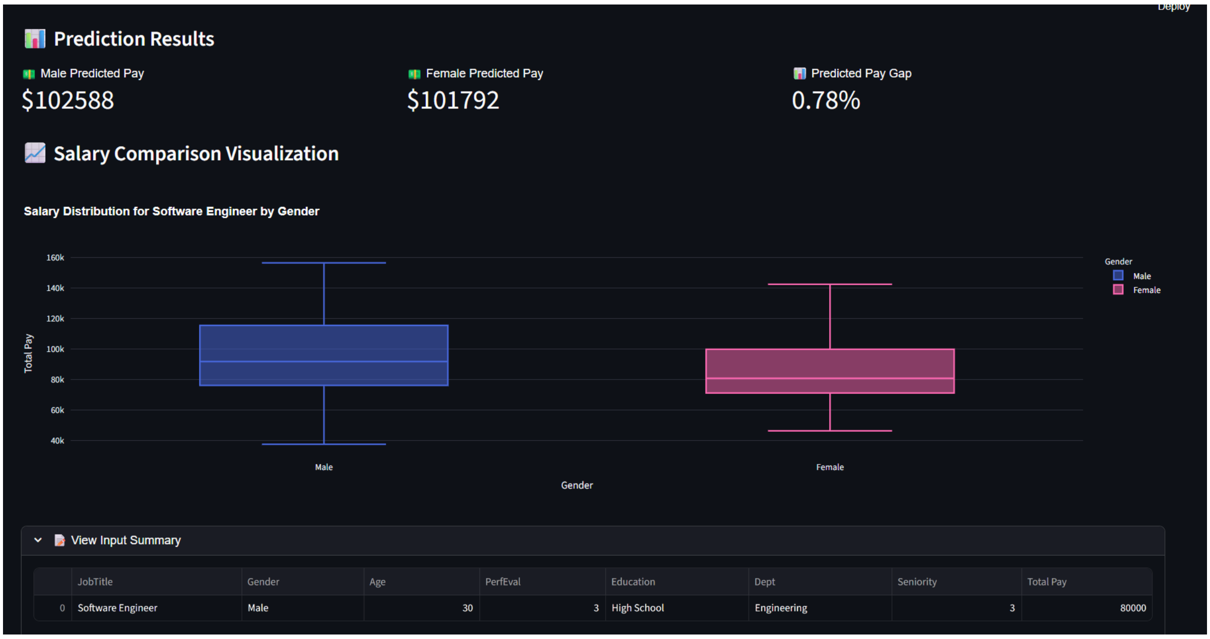
Dashboard 2 — HR Simulation (Pay-Gap Prediction)

Purpose:
Simulates potential pay outcomes by adjusting individual or role-level attributes.

How to Use:

1. Navigate to the **HR Simulation** tab.
2. Adjust any of the following parameters:
   - Base Salary
   - Job Title
   - Gender
   - Age
   - Performance Evaluation
   - Education Level
   - Department
   - Seniority
3. Results update instantly as parameters change.





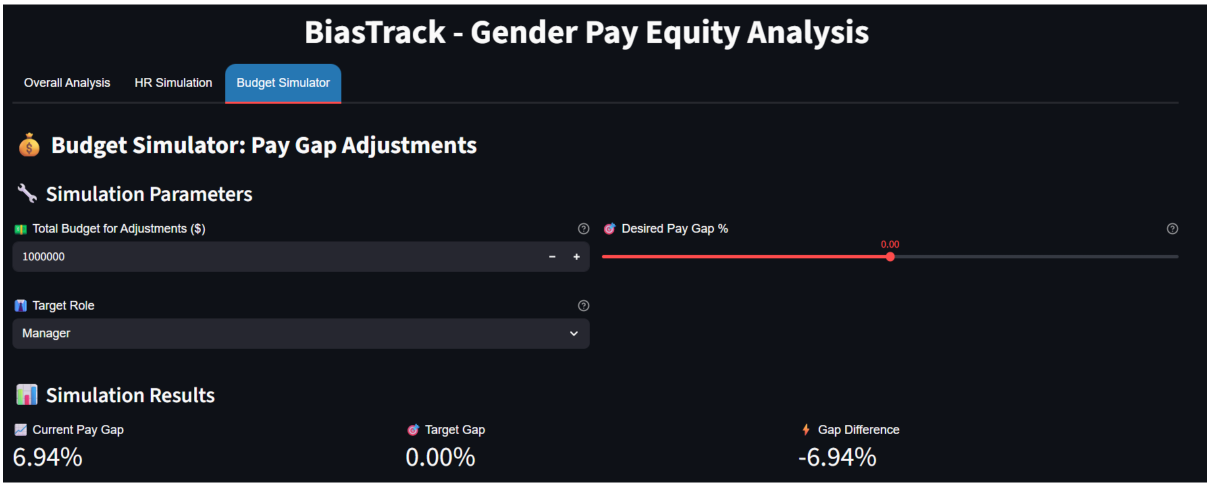
Interpretation:
Use this dashboard to test *what-if* scenarios — for example, how promotions, performance improvements, or salary revisions might affect pay equity.

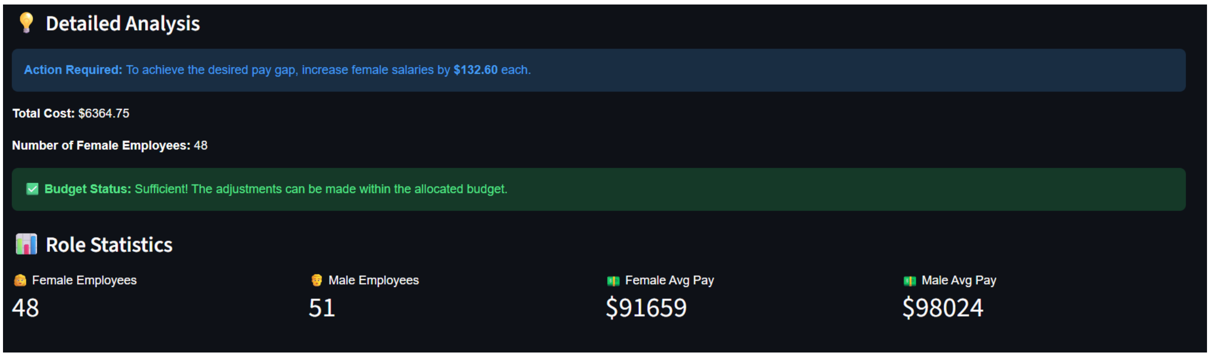
Dashboard 3 — Budget Simulator (Pay-Gap Adjustments)

Purpose:
Helps HR professionals evaluate how close they can get to pay equity within a defined budget.

How to Use:

1. Open the **Budget Simulator** tab.
2. Enter:
   - Total available budget
   - Target job role
   - Desired pay-gap percentage
3. The system automatically calculates:
   - Cost of adjustments required to close the gap
   - Whether the budget is sufficient
   - New projected salaries after adjustment





Interpretation:
Use this dashboard to plan realistic, fair salary corrections without exceeding your organisation’s financial limits.

14. Appendix: Note on the Use of AI

14.1. Purpose

This appendix provides a declaration on the use of Artificial Intelligence (AI) tools during the development and documentation of the **BiasTrack** project.

14.2. Use of AI Tools

AI assistance was used exclusively through **OpenAI's ChatGPT (GPT-5)** to support the project in the following areas:

- **Documentation Support:** Drafting and refining written materials such as the project proposal, report sections, and presentation scripts.
- **Code Structuring and Commentary:** Assisting in designing Python class structures, docstrings, and inline comments for clarity and maintainability.
- **Language Polishing:** Enhancing sentence flow, grammar, and formatting consistency across technical and business documentation.
- **Formatting and Report Styling:** Reformatting content for readability and structural alignment with NUS-ISS report guidelines.

14.3. Scope of AI Contribution

- The AI tool was used strictly as a **writing and reasoning assistant**.
- All **concepts, implementation logic, system design, model training, and dashboard development** were entirely carried out by the project authors.
- No proprietary, confidential, or sensitive organisational data was shared with the AI system at any stage.
- All technical results, evaluations, and dashboards were manually verified and validated by the project team before inclusion in the report.

14.4. Intellectual Responsibility

Final responsibility for the design, content, and accuracy of the BiasTrack system and documentation rests solely with the project authors.

The AI system served only as a tool for language generation, technical drafting, and refinement—comparable to an intelligent assistant—and not as a source of original research, programming, or analytical conclusions.

15. References

1. World Economic Forum (2025). *Global Gender Gap Report 2025.* Retrieved from https://www.weforum.org/publications/global-gender-gap-report-2025
2. Ministry of Manpower (2024). *Update on Singapore's Adjusted Gender Pay Gap.* Retrieved from https://stats.mom.gov.sg
3. Equal Pay Today (2025). *Gender Pay Gap Statistics.* Retrieved from https://www.equalpaytoday.org/gender-pay-gap-statistics
4. ADP (2025). *EU Pay Transparency Directive FAQs.* Retrieved from https://www.adp.com/spark/articles/2025/04/eu-pay-transparency-directive-faq.aspx
5. Withersworldwide (2025). *Singapore's Workplace Fairness Act and Its Impact on Women in the Workplace.* Retrieved from https://www.withersworldwide.com
6. Mercer (2024). *Achieving Equal Pay under the New EU Pay Transparency Directive.* Retrieved from https://www.mercer.com
7. Statista (2024). *Gender Pay Gap Infographics.* Retrieved from https://www.statista.com