# Page 1: *The Data Job Observatory*

## Overall Purpose

The **Data Job Observatory** is the most advanced analytical page of the project. It is designed as a **dynamic monitoring and interpretation tool** that tracks how the data job market evolves over time, rather than providing a static snapshot.

While *The Big Picture* summarizes the market at a given moment, this page focuses on **temporal dynamics, skill drift, and user positioning within the market**.

## Conceptual Role in the Project

This page acts as:

- a **longitudinal observatory** of the job market,
- a bridge between **historical trends** and **current signals**,
- an interpretive layer connecting market evolution to **individual user profiles**.

It emphasizes *change over time* rather than simple counts.

## Data Loading and Normalization

- Data is loaded from the /api/data endpoint.
- Each job posting is parsed and enriched client-side:
  - publication dates are converted into monthly time units,
  - job titles are classified into **normalized role categories** (Data Analyst, Data Scientist, Data Engineer, etc.),
  - technical skills, tools, soft skills, and domains are parsed into standardized lists,
  - salary, experience level, hybrid policy, and visa sponsorship are normalized into comparable formats.

This preprocessing step enables **temporal aggregation and cross-filter consistency**.

## Global KPIs (Market Pulse)

At the top of the page, three **real-time KPIs** summarize short-term market dynamics:

- **Total job offers over the last 30 days**, compared to the previous 30-day period,
- **Hottest role**, based on recent hiring volume,
- **Technology in the spotlight**, identified through relative growth in skill mentions.

These indicators are recalculated dynamically based on the active filters.

## Unified Filtering System

A persistent filter sidebar controls **all charts simultaneously**, ensuring analytical coherence.

Available filters include:

- full-text search (title and description),

- minimum annual salary,
- rolling time window (all data, last 30 days, last 60 days),
- job conditions (remote/hybrid, visa sponsorship),
- countries,
- normalized roles,
- seniority levels,
- core technical skills (AND logic, via interactive chips).

Filters can be combined freely, allowing users to observe how **market signals react to specific constraints**.

## Main Analytical Visualizations

**1. Job Volume Over Time**

A dual-line chart compares:

- the full dataset (baseline),
- the currently filtered subset.

This visualization highlights **market acceleration, slowdowns, or stability**, while preserving context.

**2. Role Drift — Skill Evolution by Role**

This chart tracks how **key technical skills evolve month by month** for a selected role.

It reveals:

- shifts in expected skill stacks,
- emerging or declining technologies,
- role specialization trends.

The chart updates dynamically when the role or filters change.

**3. Soft Skills Dynamics**

A temporal line chart shows the **evolution of soft skill mentions** over recent months.

This visualization emphasizes the growing importance of **non-technical competencies** and how their demand fluctuates across time.

**4. My Market Worth (MMW)**

This module allows users to simulate their **position within the dataset**:

- users select their core skills and experience level,
- the system estimates how many offers match their profile,
- a salary distribution is computed for comparable roles.

This feature reframes the dataset as a **reflective tool**, connecting aggregate trends to individual trajectories.

**5. Geographical Distribution (World Map)**

An interactive world map displays the **spatial distribution of job offers**, using proportional symbols per country.

This visualization supports:

- global comparisons,
- identification of regional hiring hubs,
- spatial filtering effects.

### Interaction Design

- All charts are interactive, with hover tooltips revealing precise counts and dates.
- Clicking on any chart opens a **modal enlargement**, enabling focused inspection.
- Animations and transitions support readability without obscuring interpretation.

### Analytical Contribution of the Page

The Data Job Observatory:

- transforms raw job postings into **temporal signals**,
- exposes **emerging trends rather than static rankings**,
- supports **self-positioning and reflexive analysis**,
- complements the project's other pages by adding **time, evolution, and projection**.

It represents the project's most interpretive layer, where data analysis, visualization, and user reflection converge.

# Page 2 : *Project Homepage / Introduction*

This page serves as the **main entry point** to the *Data Universe* project. Its role is to introduce the **concept, objectives, and scope** of the platform before users engage with the analytical tools.

It presents the project as a **data job market observatory**, combining job scraping, text analysis, clustering, and interactive visualization to make a fragmented and fast-evolving market readable. The page explains the **motivation**, outlines the **methodology and technical tasks**, highlights **key insights**, and documents **challenges encountered** during development.

Rather than delivering analysis results, this page provides the **conceptual and methodological context** that frames all subsequent pages (*The Big Picture*, *The Job Explorer*, *The Data Job Observatory*, *The Hidden Trends*), ensuring that data exploration is grounded in a clear analytical intent.

# Page 3 : *The Big Picture*

## Overall Purpose

This page provides a **high-level analytical overview of the data job market**. It is designed as an **entry point into the dataset**, offering users a "30,000-foot view" before they move toward more detailed exploration tools such as *The Job Explorer* or *The Hidden Trends.*

Its goal is to **summarize the structure, scale, and key patterns** of the job market through aggregated indicators and visualizations.

## Data Loading and Pre-processing

- All visualizations are powered by data fetched from the backend endpoint /api/stats-data.
- Upon loading, the dataset is **cleaned and standardized client-side**:
  - numeric fields (salary, experience) are converted to numbers,
  - boolean fields (hybrid policy, visa sponsorship) are normalized,
  - multi-value fields (skills, tools, domains) are parsed into lists,
  - missing or inconsistent textual fields are given fallback values.

This preprocessing step ensures **reliable aggregation and comparability** across all charts.

## Key Performance Indicators (KPIs)

At the top of the page, three **global indicators** summarize the dataset:

- **Total number of jobs analyzed**
- **Number of unique hiring companies**
- **Median annual salary (USD)**, computed only from valid annual USD salaries

These KPIs provide immediate context regarding the **size, diversity, and economic level** of the job market represented.

## Analytical Sections and Visualizations

The page is organized into thematic sections, each answering a specific analytical question.

**1. Who is Hiring & What For?**

- Bar charts showing:
  - **Top 10 job titles**
  - **Top 10 hiring companies**
- These charts highlight dominant roles and major market players.

**2. The Core Toolkit & The Reward**

- **Top technical skills** bar chart (e.g. Python, SQL, cloud tools)
- **Salary distribution histogram**, showing concentration ranges and high-end outliers
- This section links **skills demand to financial outcomes**.

**3. Who, What, and Where**

- Bar charts for:
  - most used tools,
  - top hiring countries,
  - most common business domains,
- A pie chart for **seniority levels**, grouping less frequent categories into "Other".

This section contextualizes jobs geographically, organizationally, and hierarchically.

**4. Workplace Policies**

- Pie charts illustrating:

- hybrid/remote vs on-site work,
      - visa sponsorship availability,
      - job posting sources.
  - These charts reflect **structural employment conditions**, not just technical requirements.

### Interaction and User Experience

- All charts are interactive:
      - hover tooltips display precise counts and percentages,
      - animations are used to progressively reveal data.
- Clicking on any chart opens a **modal view**, allowing the user to inspect the visualization in isolation, without leaving the page.

This interaction design supports **focused analysis while preserving narrative flow**.

### Analytical Role of the Page

*The Big Picture* page serves as:

- a **context-setting dashboard**,
- a bridge between raw data and detailed exploration,
- a narrative layer that transforms aggregated statistics into **interpretable insights**.

It complements:

- *The Job Explorer* (micro-level, offer-by-offer analysis),

and *The Hidden Trends* (latent semantic structures),
by anchoring the project in **clear, readable market-level patterns**.

# Page 4 : *The Job Explorer*

### Overall Purpose

This page functions as an **interactive exploration tool for individual job postings** within the *Data Universe* project. It allows users to **browse, filter, and analyze data-related job offers** (Data Analyst, Data Scientist, Data Engineer, etc.) using a combination of economic, geographic, and semantic criteria.

Unlike the dashboard or high-level visualization pages, this page adopts a **micro-level, job-centric perspective**, focusing on individual offers rather than aggregate trends.

### Interface Structure

The layout is based on a **two-column structure**:

- **Left column:** interactive filter panel
- **Right column:** dynamically updated list of job offers

This structure supports a progressive exploration workflow, where each user interaction immediately updates the results.

## Data Loading and Pre-processing

- Job data is retrieved from the backend API endpoint /api/stats-data.
- Each job offer is normalized on the client side:
    - salary values are converted to numeric formats,
    - multi-value text fields (technical skills, tools, domains, benefits, etc.) are parsed into lists,
    - boolean attributes (remote/hybrid policy, visa sponsorship) are standardized.

This preparation step ensures **robust filtering logic and consistent comparisons** across offers.

## Interactive Filtering System

The left-hand panel provides multiple types of filters:

**Continuous and Boolean Filters**

- Text search (job title and description)
- Minimum salary (range slider)
- Remote / hybrid policy
- Visa sponsorship

**Categorical Filters (Dynamic Checkboxes)**

- Countries (top 20)
- Seniority levels (top 6)
- Technical skills (top 15 – AND logic)
- Business domains (top 12 – OR logic)

Filter values are **computed dynamically from the dataset**, based on frequency counts. This prevents irrelevant or empty categories and keeps the interface grounded in the actual data distribution.

A global reset button allows users to clear all filters at once.

## Filtering Logic

Displayed job offers result from a **combined filtering logic**:

- some criteria act as strict constraints (AND),
- others allow partial matches (OR),
- salary and text search function as threshold filters.

Filtering is performed **in real time**, with no page reloads, ensuring smooth exploratory interaction.

## Results Display

- Job offers are presented as **individual job cards**.
- Each card includes:
    - job title,

- company name,
- location,
- salary information (when available),
- a short preview of key technical skills.

For performance and readability reasons, the display is capped at the **top 50 matching offers**.

## Job Detail Modal

Clicking on a job card opens a **modal window** displaying:

- the full job description,
- technical skills, tools, and soft skills,
- listed benefits,
- workplace policies (remote/hybrid, visa),
- a direct link to the original job posting.

This modal enables **in-depth analysis without breaking the exploration flow**.

## Analytical Objective of the Page

The Job Explorer is designed to:

- support **fine-grained, user-driven exploration of the data job market**,
- make concrete relationships visible between skills, salary levels, locations, and HR policies,
- complement the project's global visualizations with a **decision-oriented, exploratory interface**.

It acts as a **bridge between search engine functionality and analytical dashboarding**, grounding abstract trends in individual job realities.

# Page 5 : *Interactive Job Map*

This page offers an **interactive, exploratory view** of the data job market based on a semantic clustering of job offers.

Each job is represented as a point on a two-dimensional map generated using **UMAP**, where spatial proximity reflects textual similarity between job descriptions. Points are colored by **topic cluster**, allowing users to visually identify coherent job families and overlaps between roles.

The interface is structured around three complementary components:

- a **left sidebar** providing filters (text search, cluster, country, seniority),
- a **central interactive map** with zoom, pan, and tooltips,
- a **right analytics panel** displaying contextual statistics for selected clusters or individual jobs.

Selecting a cluster or a job dynamically updates the analytics panel, revealing information such as dominant technical skills, soft skills, business domains, seniority distribution, geographic concentration, and salary indicators when available. A dedicated **cluster guide** supports interpretation by documenting topic labels and representative skills.

This page constitutes the **most granular and exploratory layer** of the platform, enabling users to navigate the semantic structure of the data job market while linking individual offers to broader market patterns.