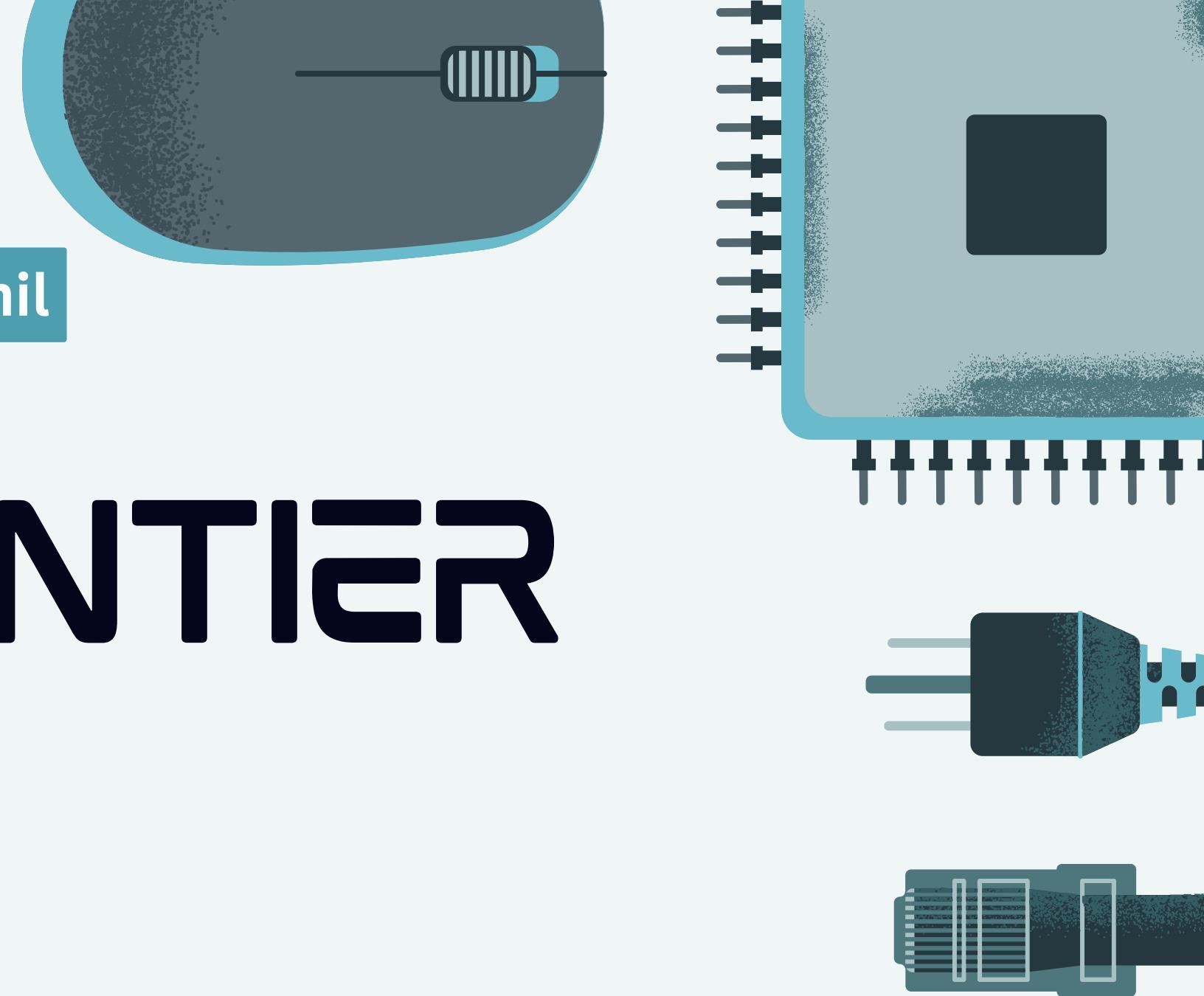
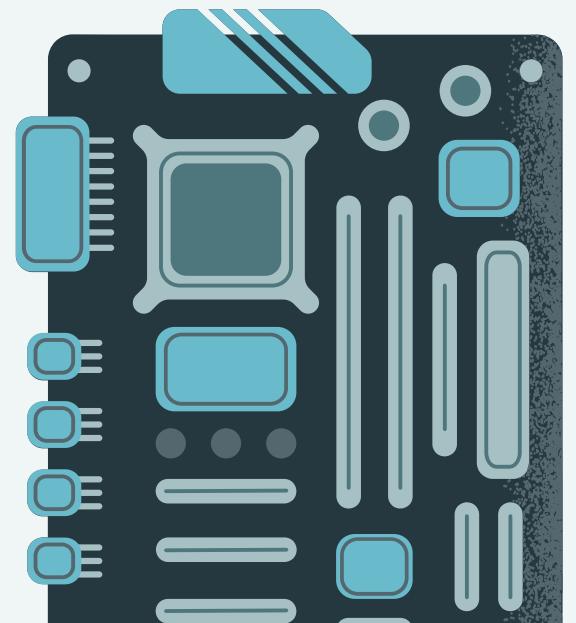
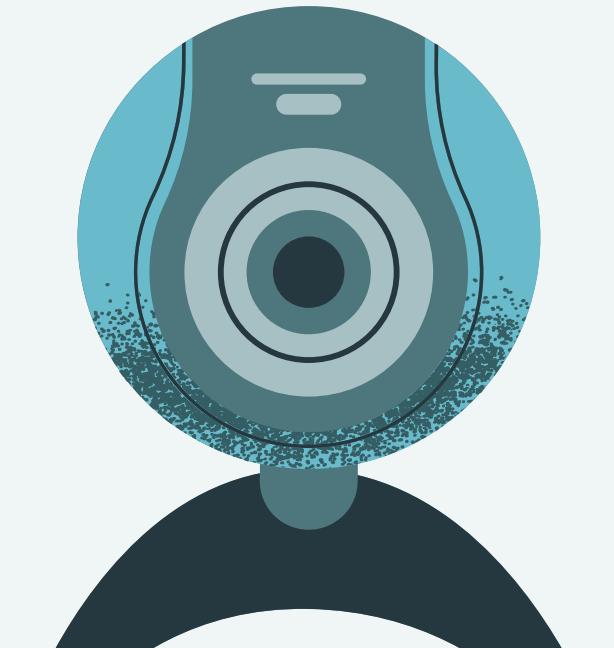
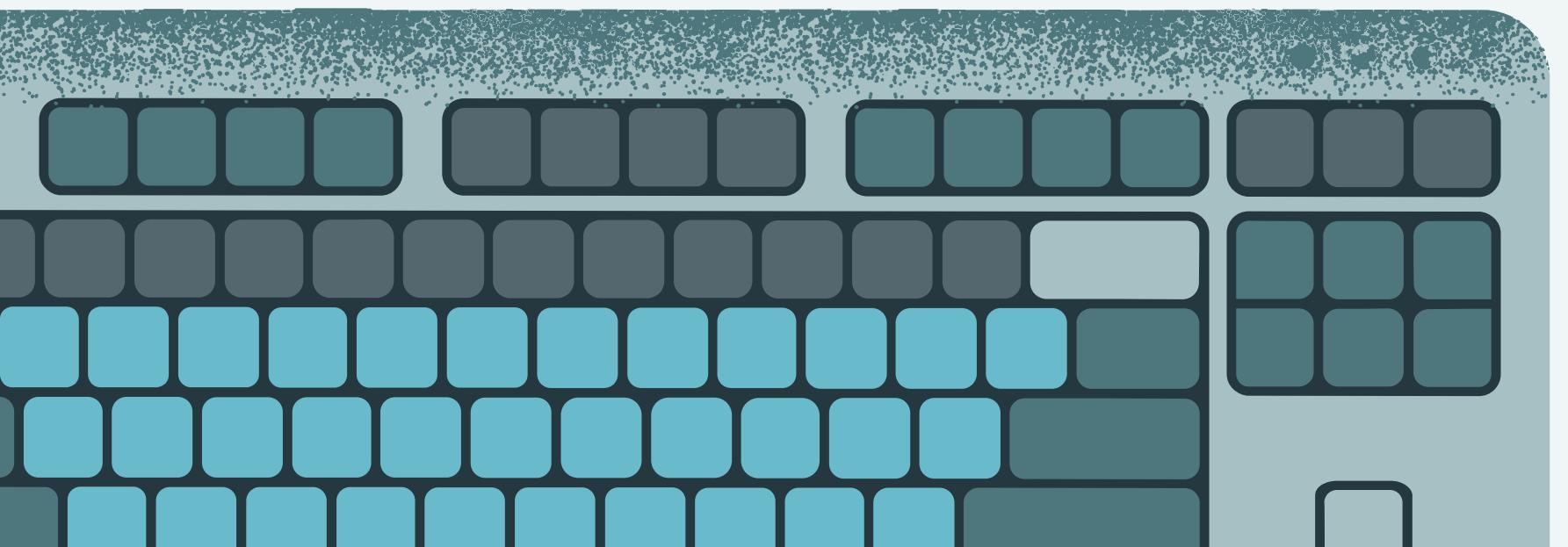


Presented by Mark, Eliska, Constanza, Aurelie & Camil

THE NEW FRONTIER IN IT HIRING



PROJECT OVERVIEW

BUSINESS CASE

We are a recruitment company dedicated to:

- Helping IT jobseekers identify potential opportunities across various fields and markets.
- Providing skilled talent resources to IT companies and market overview.

Our goal is to offer a comprehensive overview of job categories, work settings, and locations to assist job applicants in navigating career options, while also supporting IT companies in making informed, data-driven decisions based on market trends.

HYPOTHESIS

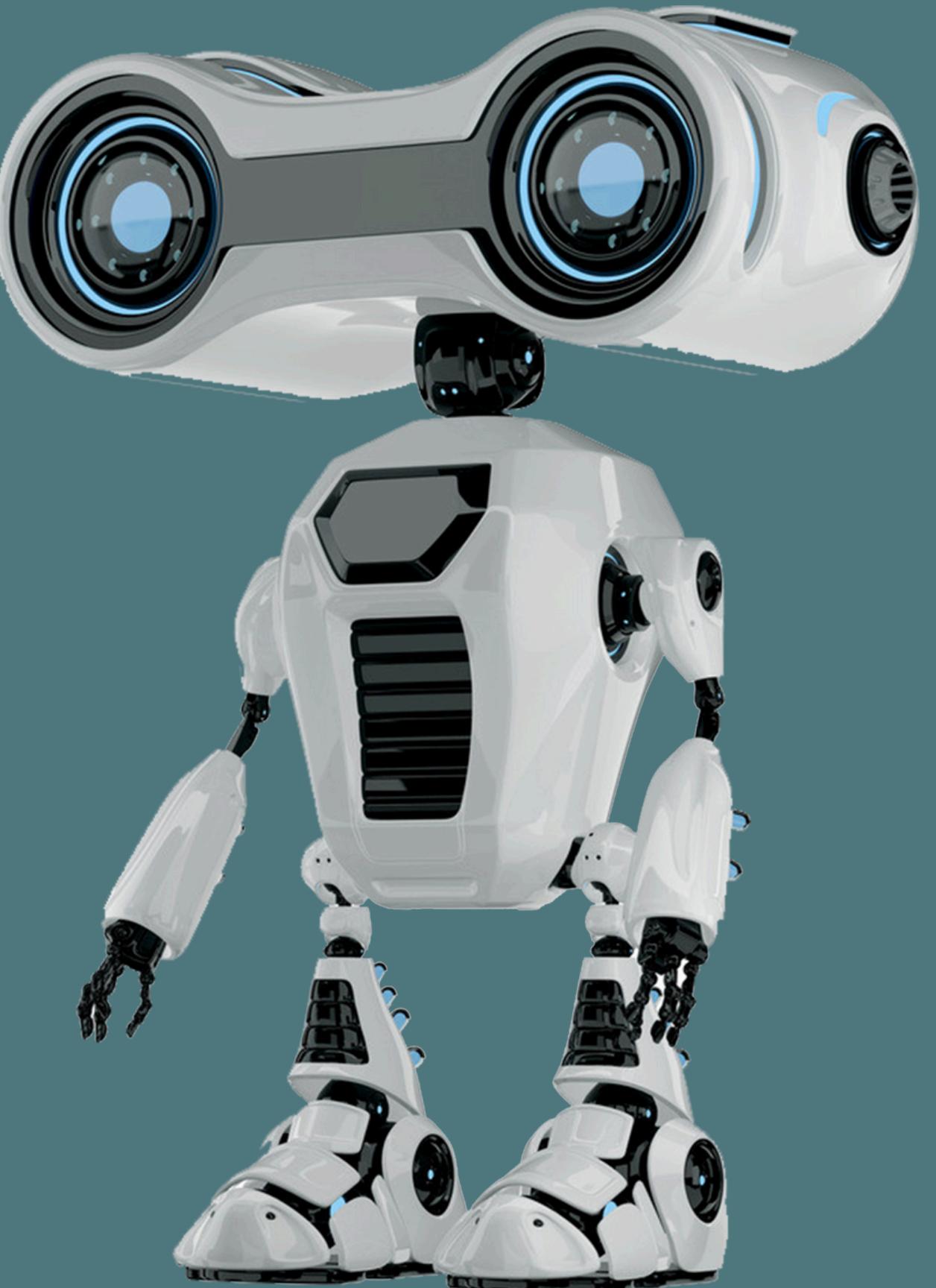
- "Jobs in specialized categories (e.g., Machine Learning, AI) offer higher salaries than general data science roles."
- "Full-time roles offer higher salaries than part-time, contract, or freelance positions."
- "Today, it may be more beneficial for companies to relocate their data teams to countries with lower salary costs or to hire employees remotely."
- "Remote positions offer higher salaries compared to in-person or hybrid roles."

DATABASE DESIGN & DATA TRANSFORMATION

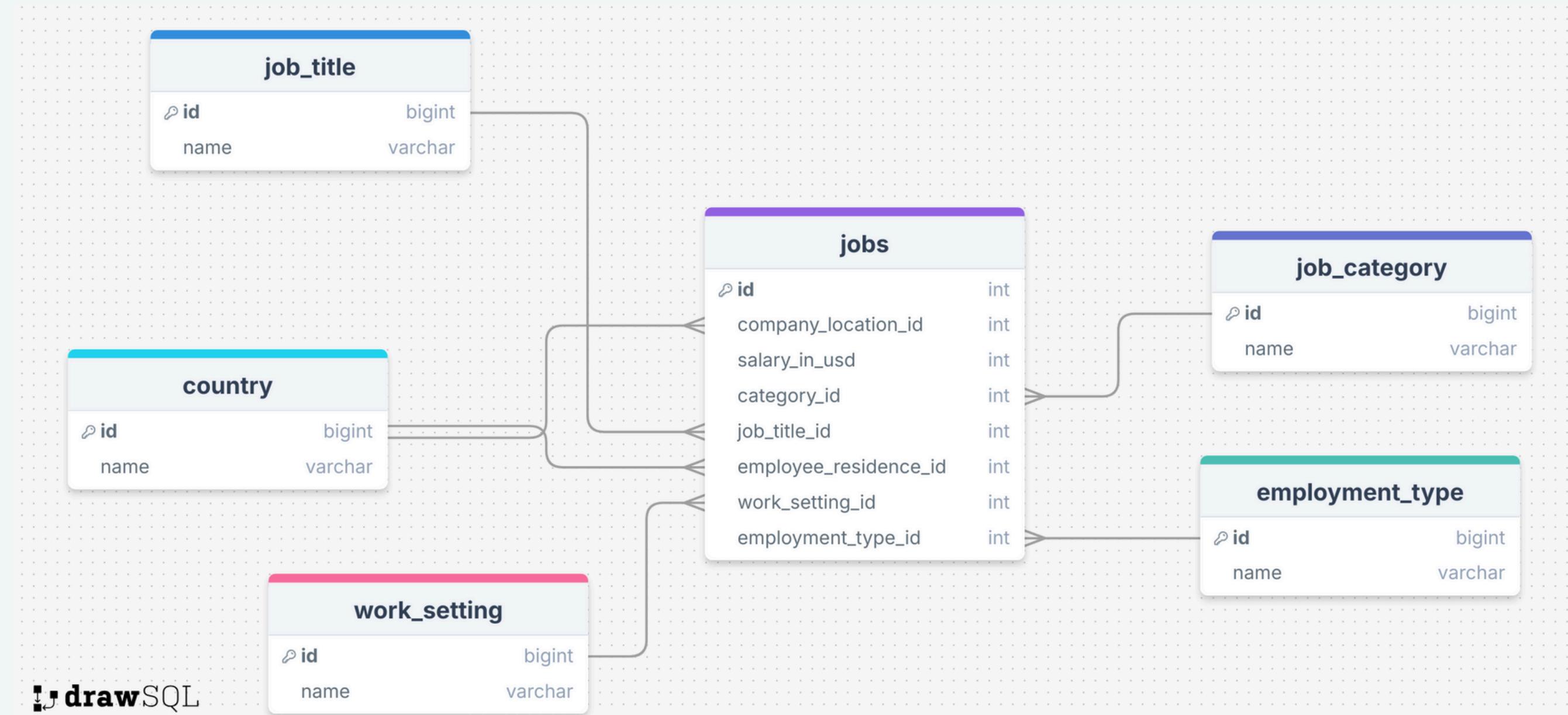
# work_year	# job_title	# job_category	# salary_currency	# salary	# salary_in_usd	# employee_reside...	# experience_level	# employment_type	# work_setting
The year when the job listing was posted or the job position is relevant (e.g., 2020, 2021).	The job title of the professional in the data science field (e.g., Data Scientist, Machine Learning Engineer).	The category of the job, specifying the main field in data science or AI it falls under (e.g., ML/AI, Data Engineering).	The currency in which the salary is provided (e.g., USD, EUR, INR).	The offered salary amount for the position, listed in the original currency.	The salary amount converted to USD for comparison purposes.	The country code where the employee resides (e.g., JP, IN, US).	The level of experience required for the job position.	employment_type:	The working arrangement for the position, specifying if the job is in-person, hybrid, or remote.
2020	Data Engineer in office	Machine Learning Engineer	EUR	19%	30.0k	JP	SE	PT	Hybrid
2022	Data Engineer	ML/AI	INR	18%	200k	IN	EX	FL	In-person
2020	Other (4273)	Other (2676)	Other (3122)	54% 62%	-198754	Other (3476)	Other (2727)	Other (2469)	Other (1621)
2022	Machine Learning Engineer in office	Analysis	EUR	186597	136086	US	MI	CT	Remote
2020	Statistician (Remote)	ML/AI	JPY	110630	67982	JP	EX	FL	Remote
2022	Machine Learning Engineer	ML/AI	INR	61280	153309	UK	MI	CT	Hybrid
2022	Data Analyst in office	ML/AI	JPY	154130	135242	DE	SE	FT	Hybrid
2020	Statistician	Data Science	EUR	172312	35156	UK	MI	FT	In-person
2020	Machine Learning Engineer	Engineering	JPY	36544	68280	CN	MI	FT	Hybrid
2022	Data Analyst in office	Data Science	JPY	178404	105324	DE	EX	PT	Remote
2021	Data Scientist (Remote)	ML/AI	JPY	187908	90706	UK	EX	CT	Remote
2022	Data Analyst			-44388	171043	UK		FL	In-person
2022	Statistician (Remote)	Engineering	us dollars	31694	73408	DE	EN	CT	Remote

DATABASE DESIGN & DATA TRANSFORMATION

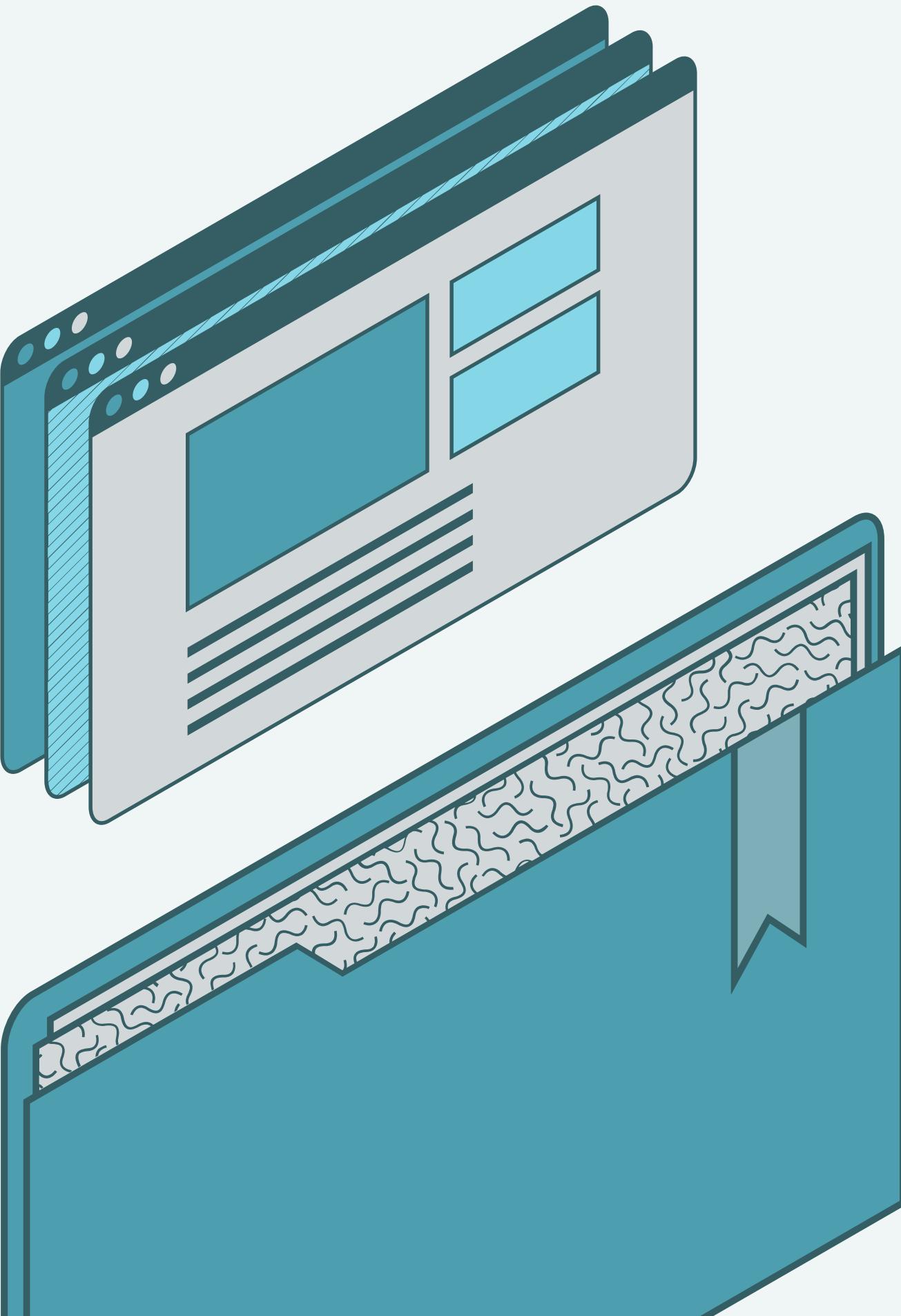
- Choose relevant **columns** based on **hypotheses**
- Created formatting **functions** only for relevant **columns**
 - Add **missing** values
 - Remove **duplicates** values
 - **Transform** values from **acronyms** to readable **text**
- Save a **CSV** file for each **SQL** table
- Creating data schemas with **drawSQL**



DATABASE DESIGN & DATA TRANSFORMATION



SQL SHOWCASE



Synthesis of the SQL request

This SQL query compares the average salaries of employees based on their residence location with the average salaries in the countries where the companies are located. It selects various fields, calculates average salaries and savings percentages. The query filters the results to include only employees whose average salary is lower than that of employees in the company's location, and sorts the results by savings percentage in descending order, highlighting the highest potential savings.

```
SELECT
    employee_country.name AS employee_residence_country, -- Employee's country of residence
    ws.name AS work_setting, -- Work setting
    company_country.name AS company_location_country, -- Company's location country
    ROUND(AVG(job.salary_in_usd)) AS avg_salary, -- Average salary in USD
    location_salary.current_location_avg_salary, -- Average salary in the company's location country
    ROUND((location_salary.current_location_avg_salary - AVG(job.salary_in_usd)) / location_salary.current_location_avg_salary * 100, 2) AS savings_percentage -- Percentage savings
FROM job
JOIN work_setting AS ws ON job.work_setting_id = ws.id -- Join with work_setting table to get work setting details
JOIN country AS employee_country ON job.employee_residence_id = employee_country.id -- then country table to get the employee's residence country
JOIN country AS company_country ON job.company_location_id = company_country.id -- and with country table to get the company's location country
JOIN (SELECT job.company_location_id,
            ROUND(AVG(job.salary_in_usd), 0) AS current_location_avg_salary -- Average salary in USD in each company location country
        FROM job
        GROUP BY job.company_location_id) AS location_salary ON job.company_location_id = location_salary.company_location_id -- Join the subquery to get average salary per company location
GROUP BY employee_residence_country, ws.name, company_location_country -- Group by employee country, work setting, and company country
HAVING avg_salary < location_salary.current_location_avg_salary -- Filter for cases where avg salary is less than the company's location average salary
ORDER BY savings_percentage DESC; -- Order by savings percentage in descending order
```

VISUALIZATION & KEY INSIGHTS

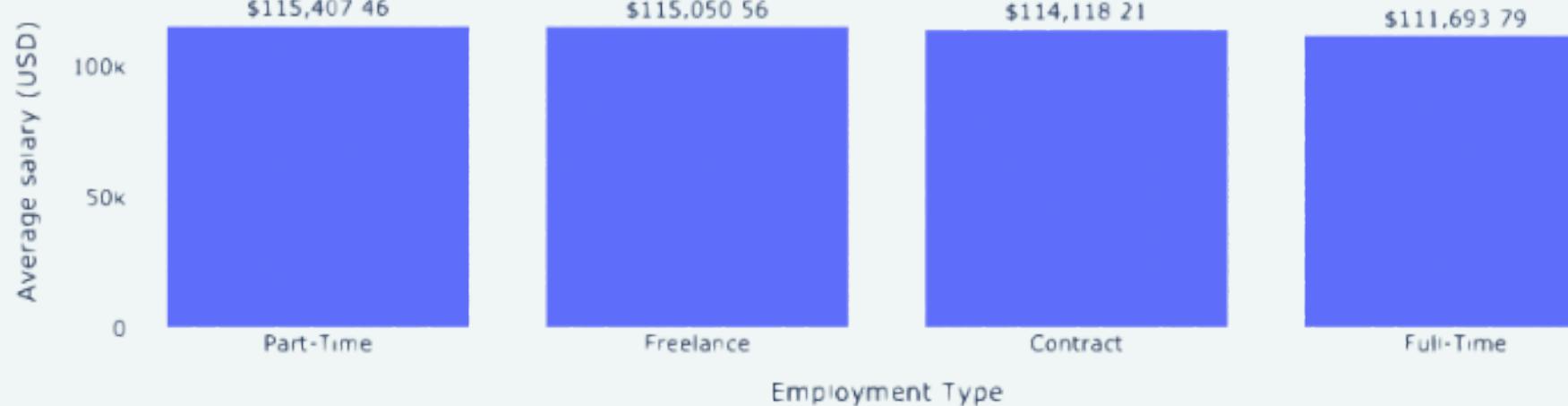
"JOBS IN SPECIALIZED CATEGORIES (E.G., MACHINE LEARNING, AI) OFFER HIGHER SALARIES THAN GENERAL DATA SCIENCE ROLES."

Average Salary by Job Title in Each Category



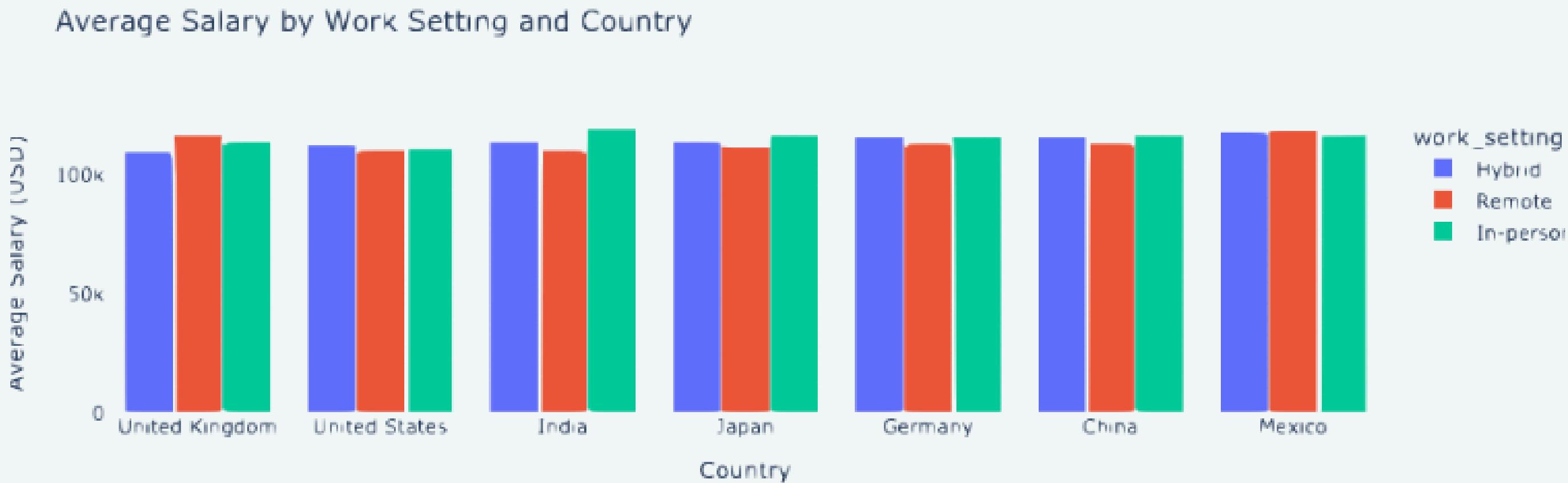
"FULL-TIME ROLES OFFER HIGHER SALARIES THAN PART-TIME, OR FREELANCE POSITIONS."

Average Salary x Employment Type

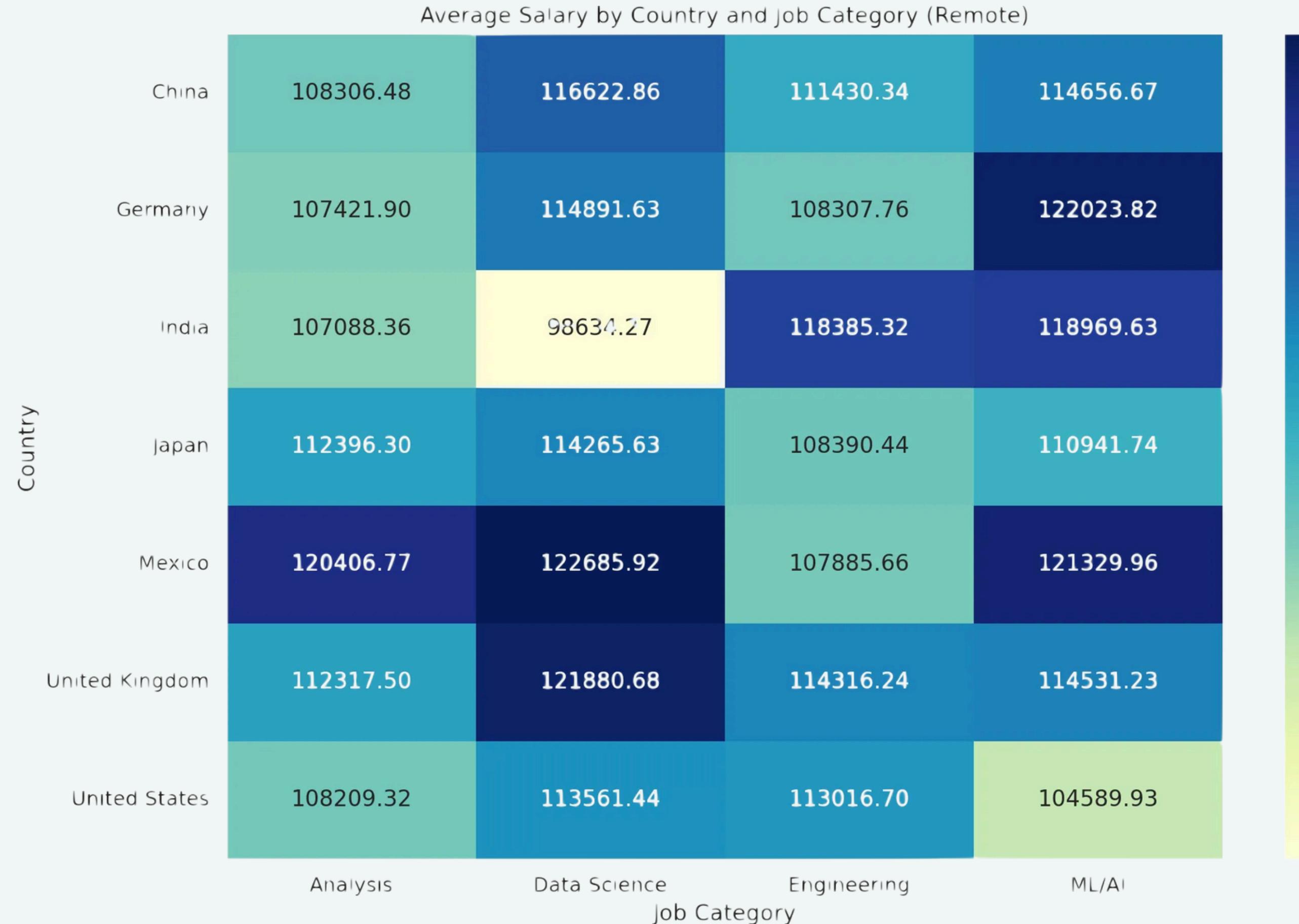


VISUALIZATION & KEY INSIGHTS

“REMOTE POSITIONS OFFER HIGHER SALARIES COMPARED TO IN-PERSON OR HYBRID ROLES.”



VISUALIZATION & KEY INSIGHTS



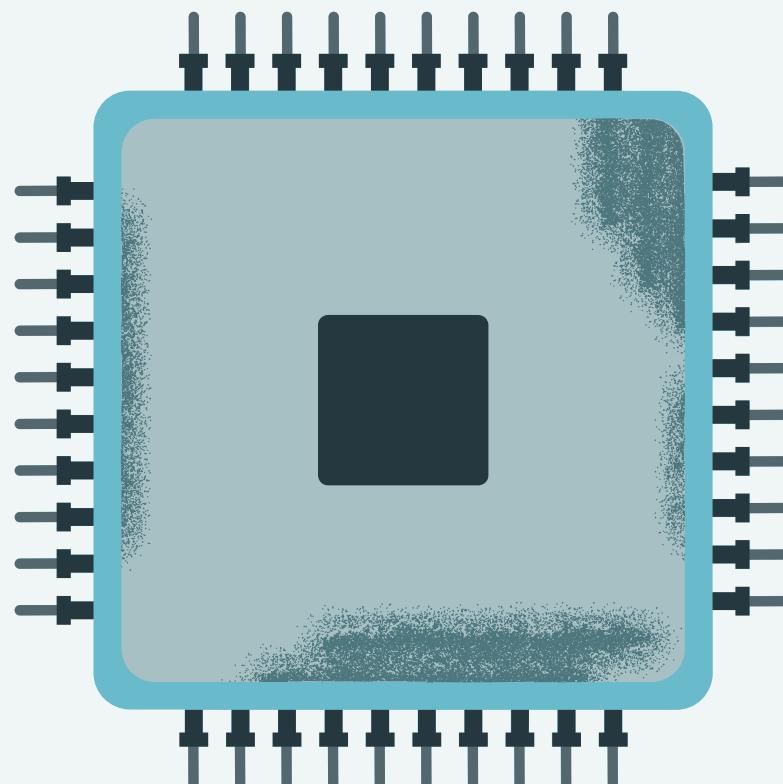
CONCLUSIONS & ADVICES

FOR JOBSEEKERS

- part-time and freelance roles often offer higher average salaries than full-time positions in IT fields
- remote roles = flexibility and access to opportunities worldwide, but in some cases (eg. India), in-person roles may offer higher salaries
- explore opportunities in countries like Japan, US, and the UK for remote Data Analyst roles

FOR IT COMPANIES

- hiring remote employees allows companies to reduce salary expenses and access a larger talent pool
- remote-friendly policies and competitive offers for remote roles to attract global talent
- understand regional variations in salary trends and work settings, while remote roles = cost savings, in-person roles in India may still be more appealing to local talent



MAJOR OBSTACLES & LEARNINGS

TRANSFERRING DATA

Obstacle: Transferring the data and tables into SQL.

Learning: breaking down the main CSV into 7 different csv's in order to create tables, using wizard import

NORMALIZATION

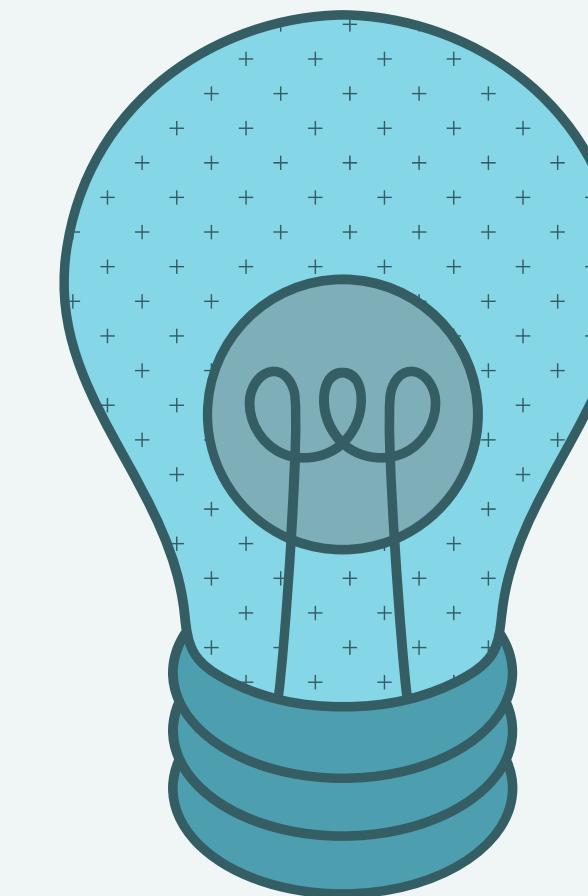
Obstacle: Break down and structure the data into related tables.

Learned how to reduce redundancy.

TABLE RELATIONSHIPS

Obstacle: Working with multiple tables.

We understood the **importance of setting accurate table** relationships to connect data effectively across tables.



ERROR DEBUGGING

We have **learned to interpret SQL error messages**, helped us troubleshoot issues faster and understand SQL syntax.

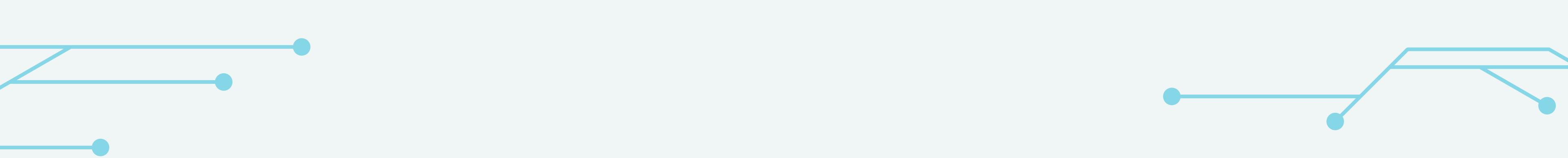
VISUALS

How to show gathered data correctly to prove our hypothesis

PROJECT MANAGEMENT

Using **Trello for task management**

Struggling in the beginning with **how to grasp our assignment**



THANK YOU!

