

Data Analysis Intern - Assignment

Objective:


Test your skills in cleaning, analyzing, and deriving insights from event-related data, just as you'd do in the internship.

Time Limit: 3 hours

Submission Format: Google Sheets + Google Doc (or PDF)

Tools Allowed: Python, Google Sheets, Excel, or any BI tool of your choice.

Dataset Provided:

Data file with >600 rows including the following columns:  Data

- `name, first_name, last_name, email, created_at`
- `approval_status, has_joined_event`
`amount, amount_tax, amount_discount, currency, ticket_name`
- `Job Title, LinkedIn profile`

Step 1: Clean the Data (30–45 min)

- Fix missing values (e.g., emails, LinkedIn URLs, job titles)
- Normalize `has_joined_event` (Yes/No) values
- Format `amount`, `amount_tax`, and `amount_discount` as numeric fields
- Extract and convert `created_at` into separate `Date` and `Time` columns

Deliverable: Cleaned version of the data as a CSV or Excel file

Step 2: Analyze Key Metrics (1–1.5 hours)

Using Excel, Google Sheets, or Python (Pandas):

1. **Conversion Funnel Metrics**
 - Total number of approved users

- Number and % who joined the event
- Show-up vs no-show rate

2. Job Role Insights

- Top 5 most common job titles
- % of students vs working professionals
- Identify any duplicates or suspicious entries

3. LinkedIn Presence

- How many profiles have a valid LinkedIn link
- How many are missing or broken (e.g., “None”, “linkedin.com/in/...”)

Deliverable: 1–2 page report or dashboard with summary tables & visualizations

Step 3: Bonus Exploration (Optional)

If you're comfortable with Python or Excel functions:

- Segment behavior based on registration time (early vs. late)
- Cluster attendees into groups: “Potential Leads”, “Engaged Attendees”, “Missing LinkedIn Info”
- Propose 2–3 **personalized re-engagement strategies** for no-shows or inactive users

Deliverable: Summary of any additional insights or re-engagement ideas

Submission Guidelines:

- Reply to the [email](#) with your Google Doc/PDF/GitHub repo link.
- Include your LinkedIn profile in the email.