Ethics for Analytics

1. Objective

The objective of this project is to understand the work culture associated with part-time jobs among students. A survey was conducted to collect relevant data, which was cleaned and prepared for analysis to ensure accuracy and usability.

2. Raw Data (File: survey1)

The raw survey data was collected online and saved in the file **survey1. This dataset contained multiple columns with survey responses.**

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	Id	Start time	Completio	Email	Name	Do you consent to participate in this survey?	What type of job have you primarily held in last 3 semesters?	How satisfied are you with the overall work environment at your part-time job?	Do you feel that your work environment promotes open communication among employees?	On a scale of 1 to 5, how supported do you feel by your supervisor or manager? (1 = Very unlikely, 5 = Very likely)	How often do you experience work-related stress at your part-time job?	Does your part-time job provide a good balance between work responsibilities and personal time?	How would you rate the physical conditions of your workplace (e.g., cleanliness, comfort, equipment)?(1 = Very unlikely, 5 = Very likely)	Rank the following factors based on how they contribute to a positive work environment for you.	Do you feel that there are opportunities for personal growth or learning in your part-time job?	How likely are you to recommend yo part-time job tc friend? (1 = Very unlike 5 = Very likely)
2	4	11/6/2024	11/6/2024	anonymous		Yes	Part-time	Neutral	Yes	4	Sometimes	Yes	4	Team collaboration;Supp	Yes	
3	5	11/6/2024	11/6/2024	anonymous		Yes	Part-time	Neutral	Yes	3	Sometimes	Yes	4	Safe and clean workspac	No	
4	6	11/6/2024	11/6/2024	anonymous		Yes	Part-time	Neutral	Yes	3	Sometimes	Yes	4	Safe and clean workspac	Yes	
5	8	11/18/202	11/18/202	anonymous		Yes	On-call	Dissatisfied	No	4	Never	Yes	3	Supportive management;	Yes	
6	9	11/18/202	11/18/202	anonymous		Yes	Part-time	Somewhat dissatisfie	Yes	4	Occasionally	No	3	Safe and clean workspac	Yes	
7	10	11/18/202	11/18/202	anonymous		Yes	Part-time	Satisfied	Yes	4	Often	Yes	5	Safe and clean workspac	Yes	
8	11	11/18/202	11/18/202	anonymous		Yes	Part-time	Very satisfied	Yes	4	Occasionally	Yes	4	Team collaboration;Safe	Yes	
9	12	11/18/202	11/18/202	anonymous		Yes	Summer Job	Satisfied	No	3	Occasionally	Yes	4	Supportive management;	Yes	
10	13	11/18/202	11/18/202	anonymous		Yes	Part-time	Satisfied	No	3	Often	Yes	4	Safe and clean workspac	No	
11	14	11/18/202	11/18/202	anonymous		Yes	Part-time	Very satisfied	No	5	Occasionally	Yes	5	Supportive management;	No	
12	15	11/18/202	11/18/202	anonymous		Yes	Part-time	Very satisfied	Yes	5	Occasionally	Yes	5	Safe and clean workspac	Yes	
13	16	11/18/202	11/18/202	anonymous		Yes	Part-time	Satisfied	No	1	Occasionally	Yes	2	Flexible hours;Team colla	No	
14	17	11/18/202	11/18/202	anonymous		Yes	Part-time	Satisfied	Yes	4	Occasionally	Yes	3	Team collaboration;Safe	No	
15	18	11/18/202	11/18/202	anonymous		Yes	Part-time	Very satisfied	Yes	5	Occasionally	Yes	5	Supportive management;	Yes	
16	19	11/18/202	11/18/202	anonymous		Yes	Part-time	Satisfied	Yes	3	Occasionally	No	4	Safe and clean workspac	Yes	

3. Data Cleaning Process:

The cleaning process was carried out in **Python** and involved the following steps:

```
import numpy as np
import pandas as pd
```

```
data = pd.read_csv(r"C:\Nivedita\Course work\Sem III\Ethics for Data Analytics\Individual Project\Survey1.csv", encoding='ISO-8859-1')
```

```
data.head()
```

	Id	Start time	Completion time	Email	Name	Do you consent to participate in this survey?	What type of job have you primarily held in last 3 semesters? \n	How satisfied are you with the overall work environment at your part-time job?	Do you feel that your work environment promotes open communication among employees?	On a scale of 1 to 5, how supported do you feel by your supervisor or manager? (1 = Very unlikely, 5 = Very likely)	How often do you experience work-related stress at your part-time job?	Does your part-time job provide a good balance between work responsibilities and personal time?	How would you rate the physical conditions of your workplace (e.g., cleanliness, comfort, equipment)? (1 = Very unlikely, 5 = Very likely)\n	Rank the following factors based on how they contribute to a positive work environment for you.	Do you feel that there are opportunities for personal growth or learning in your part-time job?	How likely are you to recommend your part-time job to a friend?\n(1 = Very unlikely, 5 = Very likely)
0	4	11/6/2024 16:09	11/6/2024 16:10	anonymous	NaN	Yes	Part-time	Neutral	Yes	4	Sometimes	Yes	4	Team collaboration;Supportive management;Safe ...	Yes	4
1	5	11/6/2024 16:19	11/6/2024 16:22	anonymous	NaN	Yes	Part-time	Neutral	Yes	3	Sometimes	Yes	4	Safe and clean workspace;Flexible hours;Team c...	No	4
2	6	11/6/2024 15:58	11/6/2024 16:22	anonymous	NaN	Yes	Part-time	Neutral	Yes	3	Sometimes	Yes	4	Safe and clean workspace;Team collaboration;Su...	Yes	4
3	8	11/18/2024 15:41	11/18/2024 15:43	anonymous	NaN	Yes	On-call	Dissatisfied	No	4	Never	Yes	3	Supportive management;Team collaboration;Safe ...	Yes	5
4	9	11/18/2024 15:59	11/18/2024 16:00	anonymous	NaN	Yes	Part-time	Somewhat dissatisfied	Yes	4	Occasionally	No	3	Safe and clean workspace;Team collaboration;Su...	Yes	3

Step 3.1: Standardizing Column Titles

The original dataset had column titles based on the survey questions.

These titles were renamed to standardized, concise formats for better readability and usability during analysis.

```
# List of new column names
new_column_names = [
    'ID',
    'StartTime',
    'CompletionTime',
    'Email',
    'Name',
    'Consent',
    'JobType',
    'Satisfaction',
    'OpenCommunication',
    'SupportBySupervisor',
    'StressFrequency',
    'WorkLifeBalance',
    'WorkplaceConditions',
    'PositiveFactors',
    'PersonalGrowth',
    'RecommendJob'
]

# Rename columns
data.columns = new_column_names
data.columns
```

```
Index(['ID', 'StartTime', 'CompletionTime', 'Email', 'Name', 'Consent',
       'JobType', 'Satisfaction', 'OpenCommunication', 'SupportBySupervisor',
       'StressFrequency', 'WorkLifeBalance', 'WorkplaceConditions',
       'PositiveFactors', 'PersonalGrowth', 'RecommendJob'],
      dtype='object')
```

Step 3.2: Dropping Unnecessary Columns

The columns Name and Email were dropped as they were empty due to the anonymous nature of the survey. This ensured adherence to ethical standards regarding participant anonymity.

Step 3.3: Transforming Ranked Options into Separate Columns

One column contained a question with four options that participants were asked to rank.

Each option was separated into its own column, named Factor 1, Factor 2, Factor 3, and Factor 4, based on the rank assigned by the participant.

```
# 1. Convert StartTime and CompletionTime to datetime
data['StartTime'] = pd.to_datetime(data['StartTime'])
data['CompletionTime'] = pd.to_datetime(data['CompletionTime'])

# 2. Standardize Yes/No Columns
yes_no_cols = ['OpenCommunication', 'WorkLifeBalance', 'PersonalGrowth']
for col in yes_no_cols:
    data[col] = data[col].map({'Yes': 1, 'No': 0})
```

```
# Remove any trailing or leading semicolons to clean the data
data['PositiveFactors'] = data['PositiveFactors'].str.strip(';')

# Split the column into multiple factors
# We use expand=True to split into separate columns
factor_columns = data['PositiveFactors'].str.split(';', expand=True)

# Rename the split columns dynamically as PositiveFactor1, PositiveFactor2, ...
factor_columns.columns = [f'PositiveFactor{i+1}' for i in range(factor_columns.shape[1])]

# Combine the new columns with the original DataFrame
data = pd.concat([data, factor_columns], axis=1)

# Drop the original PositiveFactors column if no longer needed
data = data.drop(columns=['PositiveFactors'])
```

4. Exporting the cleaned Data (File: survey2)

The cleaned dataset was saved as survey2 in Excel format for further analysis.

```
data.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 15 entries, 0 to 14
Data columns (total 17 columns):
 #   Column                Non-Null Count  Dtype
---  ------                --------------  -----
 0   ID                    15 non-null     int64
 1   StartTime             15 non-null     datetime64[ns]
 2   CompletionTime        15 non-null     datetime64[ns]
 3   Consent               15 non-null     object
 4   JobType               15 non-null     object
 5   Satisfaction          15 non-null     object
 6   OpenCommunication     15 non-null     int64
 7   SupportBySupervisor   15 non-null     int64
 8   StressFrequency       15 non-null     object
 9   WorkLifeBalance       15 non-null     int64
 10  WorkplaceConditions   15 non-null     int64
 11  PersonalGrowth        15 non-null     int64
 12  RecommendJob          15 non-null     int64
 13  PositiveFactor1       15 non-null     object
 14  PositiveFactor2       15 non-null     object
 15  PositiveFactor3       15 non-null     object
 16  PositiveFactor4       15 non-null     object
dtypes: datetime64[ns](2), int64(7), object(8)
memory usage: 2.1+ KB
```

```
# Save the DataFrame to a CSV file
data.to_csv('Survey2.csv', index=False)

# Provide a download link
from IPython.display import FileLink
FileLink('Survey2.csv')
```

Survey2.csv

File: survey2

ID	StartTime	CompletionTime	Consent	JobType	Satisfaction	OpenCommunication	SupportBy Supervisor	StressFrequency	WorkLife Balance	Workplace Conditions	Personal Growth	Recommend Job	Positive Factor1	Positive Factor2	Positive Factor3	Positive Factor4
4	2024-11-06 16:09	2024-11-06 16:10	Yes	Part-time	Neutral	1	4	Sometimes	1	4	1	4	Team col	Supporti	Safe and	Flexible hours
5	2024-11-06 16:19	2024-11-06 16:22	Yes	Part-time	Neutral	1	3	Sometimes	1	4	0	4	Safe and	Flexible	Team co	Supportive management
6	2024-11-06 15:58	2024-11-06 16:22	Yes	Part-time	Neutral	1	3	Sometimes	1	4	1	4	Safe and	Team co	Supporti	Flexible hours
8	2024-11-18 15:41	2024-11-18 15:43	Yes	On-call	Dissatisfied	0	4	Never	1	3	1	5	Supportiv	Team co	Safe and	Flexible hours
9	2024-11-18 15:59	2024-11-18 16:00	Yes	Part-time	Somewhat di	1	4	Occasionally	0	3	1	3	Safe and	Team co	Supporti	Flexible hours
10	2024-11-18 16:03	2024-11-18 16:05	Yes	Part-time	Satisfied	1	4	Often	1	5	1	5	Safe and	Team co	Supporti	Flexible hours
11	2024-11-18 16:04	2024-11-18 16:05	Yes	Part-time	Very satisfie	1	4	Occasionally	1	4	1	5	Team col	Safe and	Supporti	Flexible hours
12	2024-11-18 16:06	2024-11-18 16:08	Yes	Summer Jc	Satisfied	0	3	Occasionally	1	4	1	5	Supportiv	Team co	Flexible	Safe and clean workspace
13	2024-11-18 16:17	2024-11-18 16:19	Yes	Part-time	Satisfied	0	3	Often	1	4	0	3	Safe and	Flexible	Team co	Supportive management
14	2024-11-18 16:29	2024-11-18 16:33	Yes	Part-time	Very satisfie	0	5	Occasionally	1	5	0	5	Supportiv	Team co	Safe and	Flexible hours
15	2024-11-18 16:36	2024-11-18 16:37	Yes	Part-time	Very satisfie	1	5	Occasionally	1	5	1	5	Safe and	Team co	Supporti	Flexible hours
16	2024-11-18 17:21	2024-11-18 17:22	Yes	Part-time	Satisfied	0	1	Occasionally	1	2	0	2	Flexible h	Team co	Safe and	Supportive management
17	2024-11-18 17:23	2024-11-18 17:24	Yes	Part-time	Satisfied	1	4	Occasionally	1	3	0	3	Team col	Safe and	Supporti	Flexible hours
18	2024-11-18 20:50	2024-11-18 20:52	Yes	Part-time	Very satisfie	1	5	Occasionally	1	5	1	5	Supportiv	Flexible	Safe and	Team collaboration
19	2024-11-18 22:25	2024-11-18 22:27	Yes	Part-time	Satisfied	1	3	Occasionally	0	4	1	3	Safe and	Supporti	Team co	Flexible hours

5. Analysis of questions:

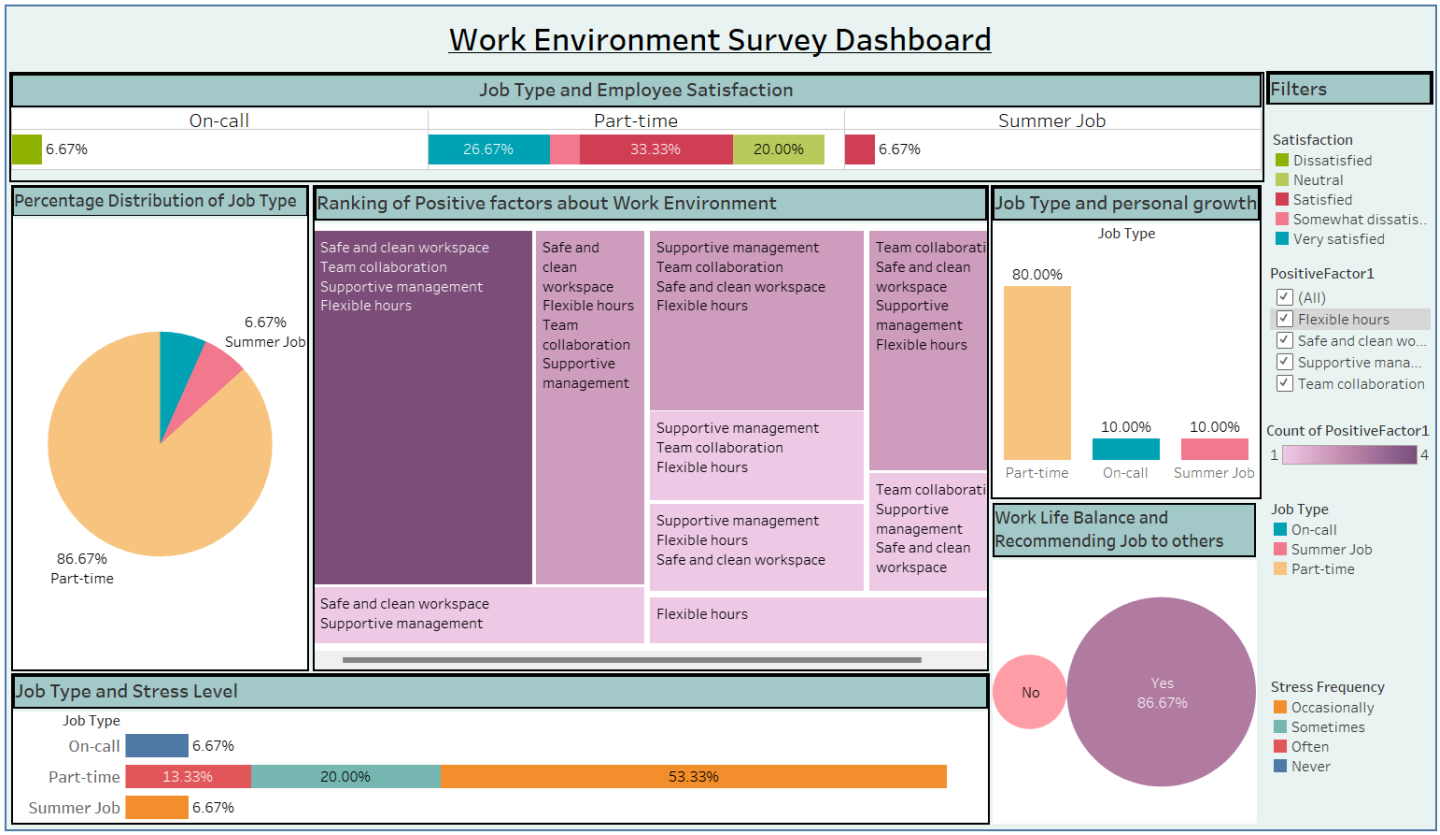
In this survey, **participants** will be answering 10 different questions about their work experiences. The questions cover various aspects of job environment, communication, management support, and work-life balance. The following types of questions are included in this form:

- **Yes / No** questions: You will choose between two options to indicate your response.
- **Ranking** questions: You will be asked to rank items in order of importance or preference.
- **Rating** questions: You will be asked to rate aspects of your work experience on a scale.
- **Multiple Choice (MCQ)** questions: You will select one options from a list of choices.

5.1 Following Research Questions addressed by this study:

1. **How does the level of stress vary across different types of part-time jobs?**
2. **What is the distribution of part-time job types among the survey participants?**
3. **What factors do employees rank as the most important positive aspects of their part-time jobs?**
4. **How does employee satisfaction differ across various job types?**
5. **What is the distribution of personal growth across different job types?**
6. **Is work life balance is deciding factor for recommending job further?**

6.Dashboard


Work Environment Survey Dashboard
Job Type and Employee Satisfaction
On-call
Part-time
Summer Job
6.67%
26.67%
33.33%
20.00%
6.67%
Filters
Satisfaction
Dissatisfied
Neutral
Satisfied
Somewhat dissatis..
Very satisfied
Percentage Distribution of Job Type
6.67%
Summer Job
86.67%
Part-time
Ranking of Positive factors about Work Environment
Safe and clean workspace
Team collaboration
Supportive management
Flexible hours
Safe and clean workspace
Flexible hours
Team collaboration
Supportive management
Supportive management
Team collaboration
Safe and clean workspace
Flexible hours
Team collaborati
Safe and clean workspace
Supportive management
Flexible hours
Supportive management
Team collaboration
Flexible hours
Team collaborati
Supportive management
Safe and clean workspace
Safe and clean workspace
Supportive management
Flexible hours
Job Type and personal growth
Job Type
80.00%
10.00%
10.00%
Part-time
On-call
Summer Job
Work Life Balance and Recommending Job to others
No
Yes
86.67%
PositiveFactor1
(All)
Flexible hours
Safe and clean wo...
Supportive mana...
Team collaboration
Count of PositiveFactor1
1
4
Job Type
On-call
Summer Job
Part-time
Job Type and Stress Level
Job Type
On-call
6.67%
Part-time
13.33%
20.00%
53.33%
Summer Job
6.67%
Stress Frequency
Occasionally
Sometimes
Often
Never