Project Phase V

Topic : Tamil Nadu Marginal Workers - Assessment

Project Definition:

Marginal Workers :

Marginal workers, often referred to as "marginal labourers," are a specific category of the labour force within a country's population. They represent a group of individuals who are engaged in irregular or low-income employment and often face economic and social vulnerability. The concept of marginal workers is commonly used in labour market and employment statistics to categorise a segment of the workforce that is distinct from regular wage/salary workers.

Marginal workers are increasing in tamilnadu as they form a part of the labour force and they work on very low wages which lead to their exploitation, the various industries where marginal workers perform include: Agriculture, Forestry and Fishing; Mining and Quarrying; Manufacturing; Electricity, Gas, steam and Air conditioning Supply; Water Supply; (Sewerage, Waste Management and remediation activities); Construction; Wholesale and Retail Trade (Repair of motor vehicles and motorcycles); Transportation and Storage; Accommodation and food service activities; Information and Communication; Financial and Insurance activities; Real Estate activities; Professional, Scientific and Technical activities; Administrative and support service activities; Public Administration and Defence, Compulsory Social Security; Education; Human Health and Social Work activities; Arts, Entertainment and recreation; Other Service Activities; Activities of Households as Employers: Undifferentiated Goods and Services; Activities of Extraterritorial Organisations and Bodies.

Marginal workers work on various types of jobs which include primary workers fields like Agriculture, Fishing, Mining and to white collar jobs like Public Administration and Defence, Scientific and Technical activities. The workers may also work in both urban or rural area, Their area of work and work type is dependent on their educational qualification which in turn determine their socio-economic status.

The age distribution of marginal workers can vary from one region or country to another, and it can change over time due to economic, social, and

demographic factors. In the context of India, where the concept of marginal workers is commonly used, the age distribution of marginal workers may typically be characterised as follows:

Child Labour: Child labour involves the employment of children below the legal working age, which varies from country to country but typically includes those under the age of 14. Child labour is a critical issue among marginal workers in many parts of the world, including India. These children often work in hazardous and exploitative conditions, and their age distribution can encompass a range of ages, from very young children to teenagers.

Younger Age Groups: Marginal workers often include a significant proportion of younger individuals, particularly those in the age group of 15-34 years. This is due to several reasons, including entry-level employment opportunities in various sectors, limited access to formal education, and the need for income generation in lower-income households.

Prime Working-Age Population: The prime working-age group, which typically spans from 35 to 59 years, also contributes to the age distribution of marginal workers. People in this age group may become marginal workers due to underemployment, seasonal employment, or limited job opportunities, especially in the informal sector.

Older Workers: While the majority of marginal workers are relatively younger, there are older individuals, typically aged 60 and above, who engage in marginal employment. This may be due to factors like a lack of retirement benefits, limited pension options, or the need to support themselves and their families.

Gender distribution among marginal workers varies depending on cultural, societal, and economic factors, as well as the specific region and industry in question. Here are some general trends in the gender distribution of marginal workers:

Gender Disparities: Gender disparities are often observed among marginal workers. In many parts of the world, including India, you may find more men than women engaged in marginal employment.

Men as Predominant Marginal Workers: In certain sectors like agriculture, construction, and transportation, men tend to make up a significant portion of the marginal worker population. They often work as agricultural laborers, construction workers, rickshaw pullers, or in other physically demanding roles.

Women in Informal and Unpaid Work: Women are more likely to be engaged in informal, part-time, or unpaid work, such as household and care work. These roles may not be recognized as formal employment, but they are essential to the functioning of households and communities.

Let us look at the problems of marginal workers and solution to be offered with application of innovation.

1. Lack of Access to Social Services: Many marginal workers struggle to access essential social services such as healthcare and clean water. Marginal workers often struggle to access healthcare services due to financial constraints or limited healthcare infrastructure in their areas. This can lead to unmet medical needs and poorer health outcomes.Adequate access to clean water and sanitation facilities is often lacking in their living conditions, posing health risks.

Innovation:

Telemedicine and eHealth Platforms:
Develop telemedicine and eHealth platforms that allow marginalized workers to consult with healthcare professionals remotely. These platforms can provide diagnosis, prescription, and follow-up care through video calls and messaging apps.

Mobile Health Apps:
Create mobile health apps that provide healthcare information, appointment scheduling, and medication reminders. These apps can also offer access to a database of common health conditions and treatments.

Electronic Health Records (EHR):
Implement EHR systems in healthcare facilities serving marginalized communities to improve patient record management and enable quick access to medical history.

Geographic Information System (GIS) for Healthcare:
Use GIS to map healthcare facilities and track patient needs in real-time. This can help optimize the allocation of healthcare resources to underserved areas.

Learning and Health Literacy:

Develop eLearning platforms and mobile apps that educate marginalized workers about health, sanitation, and disease prevention. This can help improve their health literacy and empower them to make informed decisions.

2. Gender Disparity: Gender disparity in the job market, particularly among marginal workers, refers to the unequal treatment and opportunities that women and men experience in their employment. This issue is characterised by differences in wages, job roles, and access to job opportunities.

Women who are marginal workers often earn less than their male counterparts for the same or similar work. This wage gap is a significant issue as it perpetuates economic inequality. Gender stereotypes often lead to occupational segregation, with women more likely to work in lower-paying, less-skilled jobs. Men, on the other hand, tend to dominate higher-paying positions.Women in marginalised communities may have fewer job opportunities available to them due to social and cultural norms that restrict their mobility and choices.

Innovation:

Digital Skills Training Programs:

Implement training programs that specifically target women in marginalized communities to develop digital skills and IT knowledge. These programs should be designed to bridge the digital gender divide and promote equal access to IT opportunities.

Online Job Platforms:

Create online job platforms that connect women in marginalized communities with IT job opportunities. These platforms should be user-friendly, promote job postings without gender bias, and offer resources for skill development.

Remote Work and Telecommuting:

Promote remote work opportunities in the IT sector, allowing women in marginalized communities to work from home and overcome mobility restrictions. This flexibility can enable more women to access IT jobs.

Diversity and Inclusion Initiatives:

Encourage IT companies to implement diversity and inclusion initiatives, including gender quotas, mentorship programs, and bias training, to create a more welcoming and equitable work environment.

3. Educational Qualification: Marginal Worker is increasing every year due to lack of proper education which makes them work physically demanding jobs and illiterate jobs which are very daunting and give in return low wage. Marginal workers often lack the necessary skills and qualifications to access jobs that offer higher wages and job security. They may be stuck in low-paying, unskilled jobs. Many marginal workers have low levels of literacy and numeracy, which restrict their ability to perform even basic job tasks or access educational resources.

Innovation:

Digital Literacy and Basic IT Training:
Launch digital literacy and basic IT training programs aimed at marginalized workers. These programs can provide foundational skills necessary for accessing online resources and starting a career in IT.

Online Learning Platforms:
Create accessible and user-friendly online learning platforms that offer IT courses, coding tutorials, and other relevant skills training. These platforms can be used by marginal workers to upskill at their own pace.

Open Educational Resources (OER):
Promote the use of open educational resources, including free online courses and textbooks, to make education more affordable and accessible to marginal workers looking to improve their skills.

Community Tech Centers:
Establish community technology centers equipped with computers and internet access, where marginal workers can access educational resources, receive training, and seek guidance from mentors.

Mobile Learning Apps:
Develop mobile learning apps that provide bite-sized, interactive lessons on IT skills and job preparation. These apps can be used on smartphones, which are more accessible to many marginalized workers.

4. Child Labour: Child Labour include working of children from the age of 5-14. This age is the age of children to study in school but they are doing physically demanding jobs which may affect physical and mental development of children. Children from marginal worker families are often forced into work at a young age, sometimes engaging in hazardous or exploitative labour to contribute to their family's

income. Child labour deprives these children of their right to education, which can have long-term consequences on their future employability and overall well-being. Child labour can perpetuate the cycle of poverty. When children are forced to work instead of attending school, they are more likely to continue in low-paying, unskilled jobs as adults.

Innovation:

Online Child Labor Reporting Portals:

Create user-friendly online platforms where concerned citizens, teachers, and community members can report instances of child labor. This can help authorities identify and intervene in cases of child exploitation.

Child Labor Monitoring Apps:

Develop mobile apps that enable labor inspectors and NGOs to collect real-time data on child labor cases. These apps can use GPS to pinpoint hotspots and facilitate targeted intervention.

Data Analytics and AI for Predictive Analysis:

Utilize data analytics and artificial intelligence (AI) to analyze data on vulnerable areas and identify patterns of child labor. Predictive analysis can help focus resources on high-risk regions.

Digital Identity and Birth Registration:

Implement digital identity and birth registration systems to ensure every child has an official record, making it easier to track school attendance and identify child labor cases.

Child-Friendly Educational Platforms:

Develop online educational platforms tailored for children who have missed out on traditional schooling due to child labor. These platforms can offer age-appropriate lessons and activities.

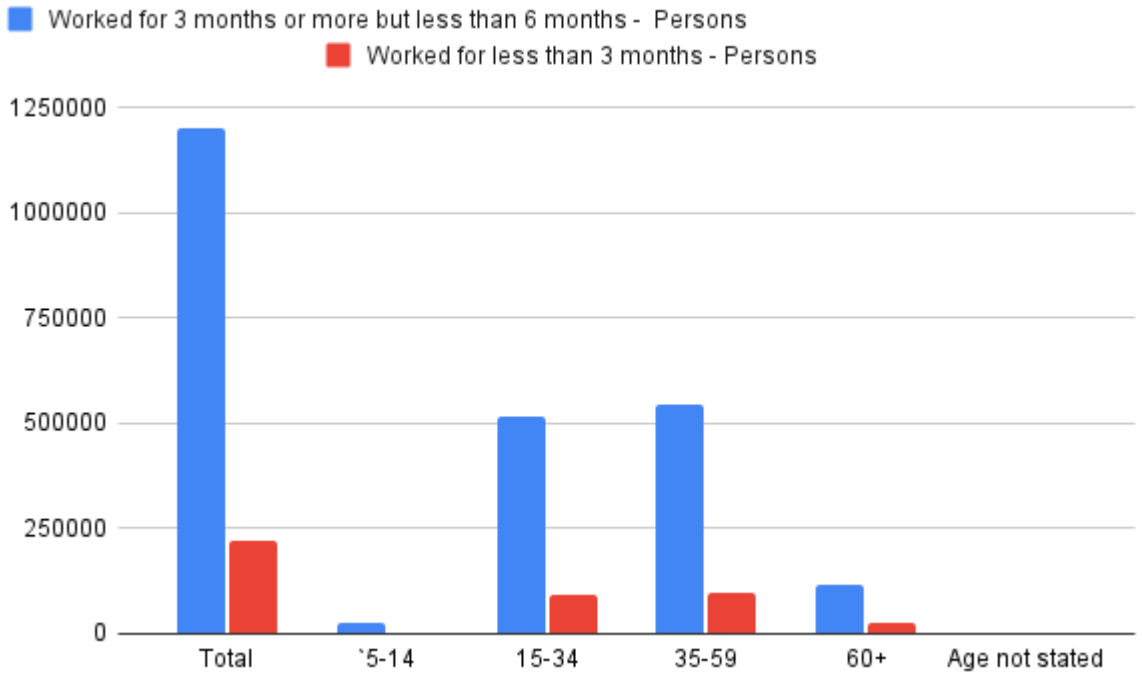
eLearning Centers:

Set up eLearning centers in areas with a high incidence of child labor, providing access to computers and the internet. These centers can offer catch-up education to bridge the learning gap.

Let us look at the problems of marginal workers and solution to be offered with reference to graphical analysis

1. Lack of Access to Social Services: Many marginal workers struggle to access essential social services such as healthcare and clean water. Marginal workers often struggle to access healthcare services due to financial constraints or limited healthcare infrastructure in their areas. This can lead to unmet medical needs and poorer health outcomes.Adequate access to clean water and sanitation facilities is often lacking in their living conditions, posing health risks.

Marginal Work job duration depend on their age and duration of work, This criteria would determine the way to provide social security to the people



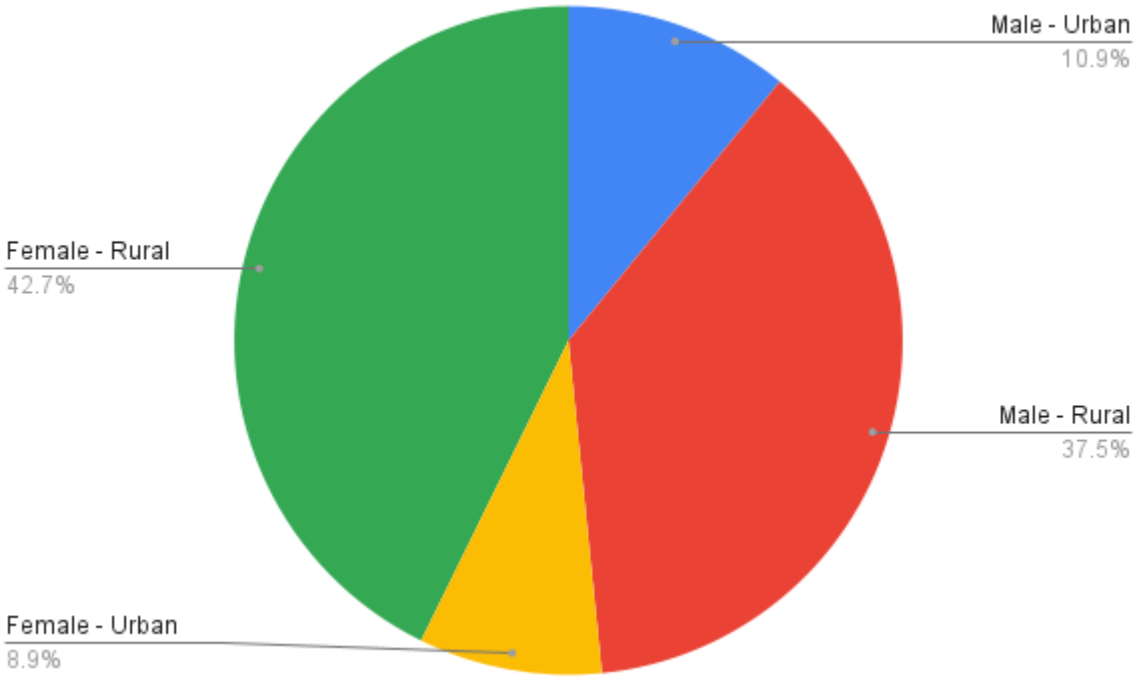
Solution: Establish mobile health clinics that reach remote areas where marginal workers live and work, providing basic healthcare services and health education.Implement sanitation programs that provide clean water sources for marginal worker living places, build toilets, and promote hygiene practices to improve living conditions and reduce health risks.

2. Gender Disparity: Gender disparity in the job market, particularly among marginal workers, refers to the unequal treatment and opportunities that women and men experience in their employment. This issue is characterised by differences in wages, job roles, and access to job opportunities.

Women who are marginal workers often earn less than their male counterparts for the same or similar work. This wage gap is a significant issue as it perpetuates economic inequality. Gender stereotypes often lead to occupational segregation, with women

more likely to work in lower-paying, less-skilled jobs. Men, on the other hand, tend to dominate higher-paying positions.Women in marginalised communities may have fewer job opportunities available to them due to social and cultural norms that restrict their mobility and choices.

Let us look at the job opportunities offered to men and women in rural area



Women are offered more jobs in rural areas but it's often non technical and physical labour. They have huge pay differences as they work for a very low wage than men. Women job contribution is less in urban area than men.

Solution: Enforce laws and regulations that ensure equal pay for equal work. Encourage companies to perform regular pay audits to identify and rectify gender-based wage gaps. Conduct awareness and training programs to sensitise employers and communities to gender issues, fostering a more inclusive and equitable work environment. Implement affirmative action policies that encourage employers to hire and promote women in all job sectors. These policies may include quotas for female representation in leadership positions.

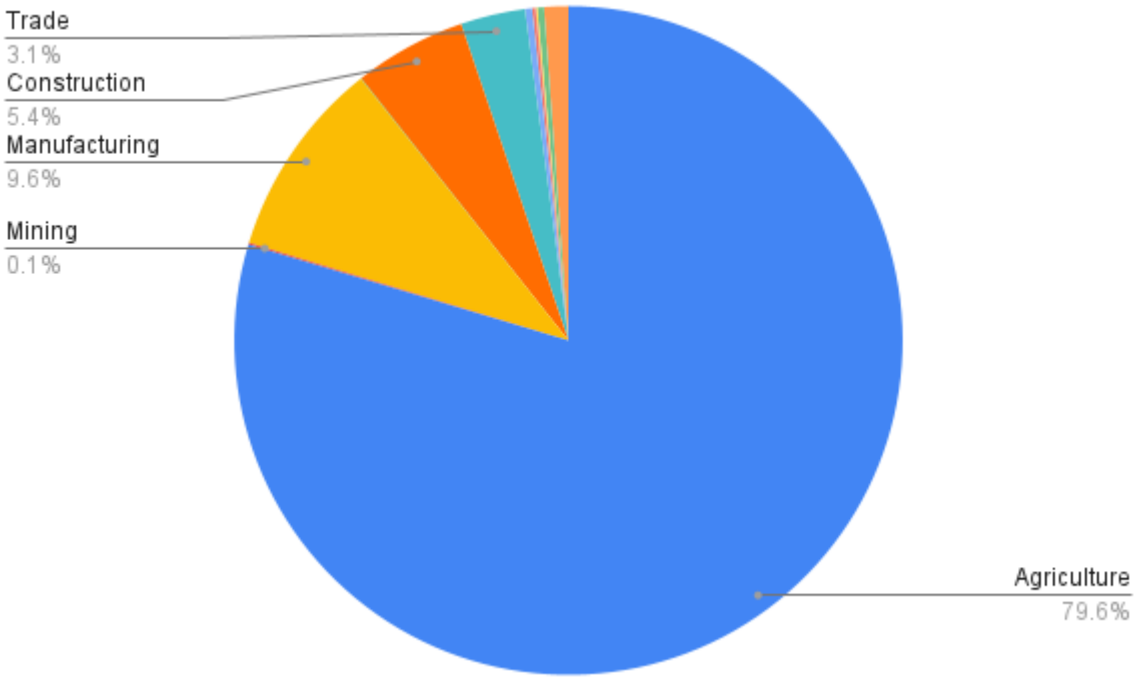
3. Educational Qualification: Marginal Worker is increasing every year due to lack of proper education which makes them work physically demanding jobs and illiterate jobs which are very daunting and give in return low wage. Marginal workers often

lack the necessary skills and qualifications to access jobs that offer higher wages and job security. They may be stuck in low-paying, unskilled jobs. Many marginal workers have low levels of literacy and numeracy, which restrict their ability to perform even basic job tasks or access educational resources.

Solution: Establish adult education and literacy programs that cater to the specific needs of marginal workers, offering flexible schedules and community-based learning.Provide vocational training and skill development programs that equip marginal workers with practical skills in demand in the job market. Offer scholarships, grants, or low-interest loans to help marginal workers, particularly women, access education and training.

4. Child Labour: Child Labour include working of children from the age of 5-14. This age is the age of children to study in school but they are doing physically demanding jobs which may affect physical and mental development of children. Children from marginal worker families are often forced into work at a young age, sometimes engaging in hazardous or exploitative labour to contribute to their family's income. Child labour deprives these children of their right to education, which can have long-term consequences on their future employability and overall well-being. Child labour can perpetuate the cycle of poverty. When children are forced to work instead of attending school, they are more likely to continue in low-paying, unskilled jobs as adults.

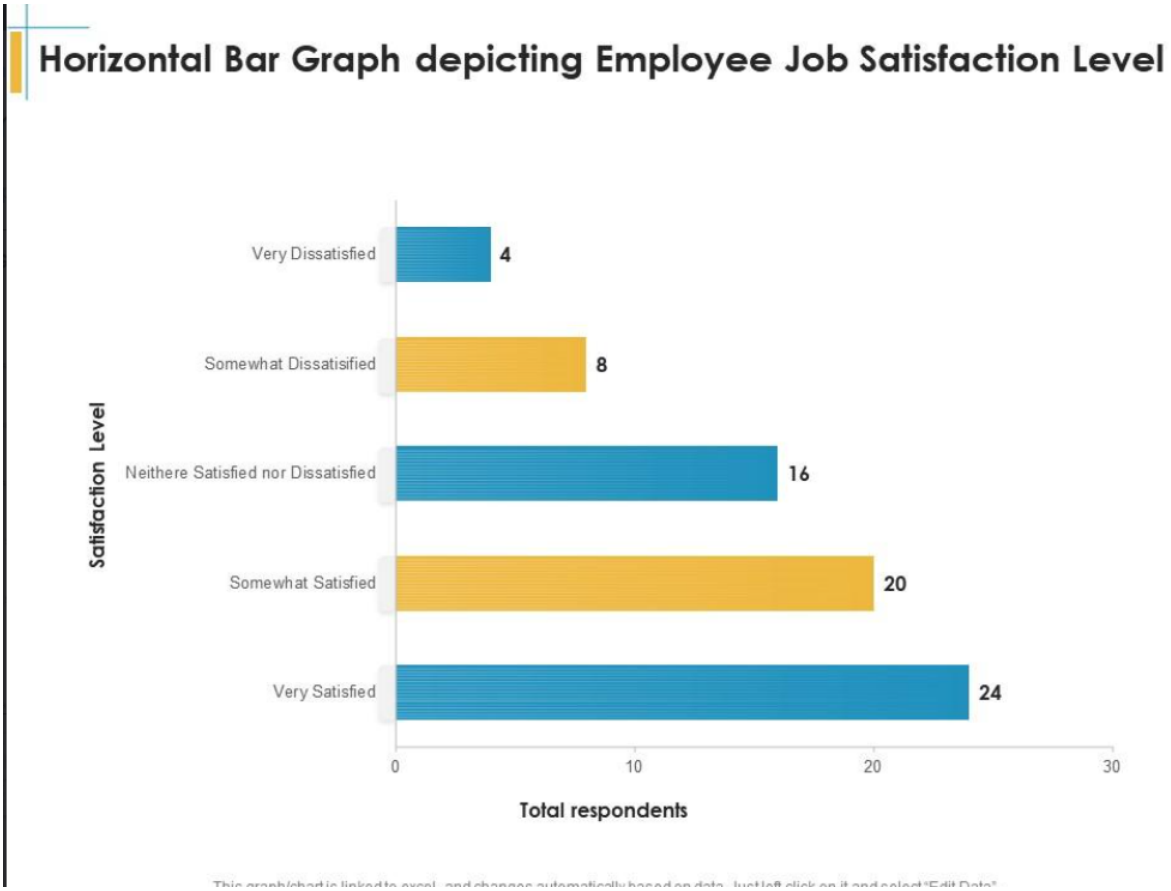
Let us look at the various job distribution of young people from 5-14



Children usually work in agriculture which is physical work which is demanding and does not require much intellect. It's very demanding and affects the overall development of children.

Solution: Child labour can be avoided by providing quality education to children which lead to increased economic status in future. Child labour can be avoided as follows, Improve access to quality education for children of marginal workers, ensuring schools are affordable, safe, and located within reasonable proximity. Enhance existing child labour laws and regulations and ensure their strict enforcement to prevent children from engaging in harmful work.

5. Job Satisfaction based on work duration: Employee job satisfaction can significantly vary based on the duration of work, particularly among migrant workers. Those who have been employed for less than 6 months may experience lower job satisfaction due to the challenges of adapting to a new environment and job role. However, as they gain more experience and familiarity with their work, those who have been employed for more than 6 months tend to see an improvement in job satisfaction, as they become more settled and skilled in their roles. These improvements can be attributed to the sense of stability and familiarity that develops over time. Migrant workers, in particular, may experience increased job satisfaction as they adapt to their new work environment and overcome initial hurdles. Therefore, the duration of work plays a crucial role in determining the job satisfaction of employees, especially migrant workers, with an upward trend in satisfaction as the duration of employment increases.

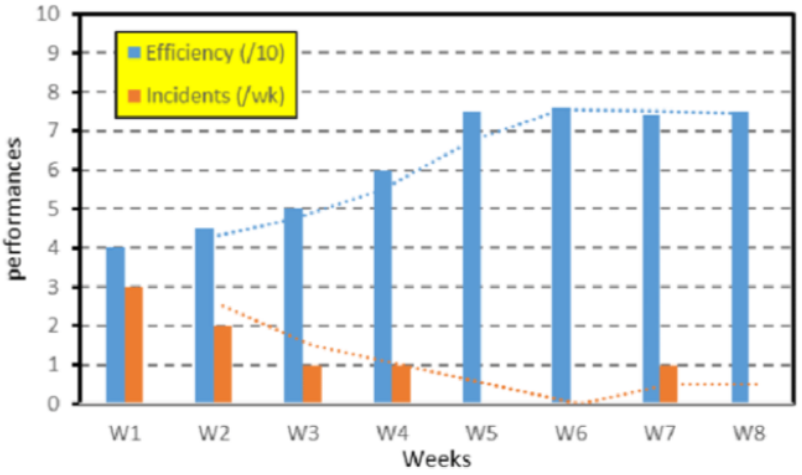


Solution: To enhance the job satisfaction of migrant workers, it is essential to focus on several key strategies. Firstly, providing comprehensive orientation and training programs can help workers adapt to their new roles and surroundings more effectively. Secondly, fostering a supportive work environment with clear communication channels and opportunities for feedback can address their unique needs and concerns. Additionally, offering benefits like health insurance, access to affordable housing, and support for their families can relieve some of the stress associated with migration. Lastly, recognizing and rewarding their contributions, both financially and through acknowledgment, can motivate and boost their job satisfaction. By implementing these strategies, employers can create a more welcoming and conducive work environment for migrant workers, ultimately leading to higher levels of job satisfaction.

6. Employee performance analysis: Analyzing employee performance based on the duration of work, particularly for migrant workers, reveals interesting insights. In the initial 3 months of employment, performance among migrant workers might be slightly below optimal due to the learning curve and adaptation to the new work environment. Those who have worked for less than 6 months may still be in the process of acclimatizing, which can affect their productivity. However, as they

accumulate more than 6 months of work experience, performance tends to improve as they become more proficient in their roles and accustomed to the workplace. Migrant workers' performance may demonstrate an upward trajectory as they overcome initial challenges and gain a better understanding of job requirements. Hence, taking into account the duration of work is essential when evaluating and managing the performance of migrant workers, with a focus on providing support and resources during their initial months to maximize their long-term contributions.

Let us look at the performance of employee across various timestamp



Sample analysis and visualize results for the employee performance analysis.

Solution: To enhance the performance of migrant workers, it is crucial to implement a multi-faceted approach. Firstly, comprehensive onboarding and training programs tailored to their needs can accelerate their adaptation to the new work environment. Secondly, creating a supportive workplace culture that promotes inclusivity, open communication, and a sense of belonging can motivate and engage migrant workers. Additionally, providing access to resources such as language courses, cultural sensitivity training, and mentorship programs can further facilitate their integration. Regular feedback and performance evaluations should be conducted to track their progress and provide constructive guidance. Finally, offering incentives, both financial and non-monetary, for high performance can boost motivation and commitment. By combining these strategies, employers can improve the performance of migrant workers, ultimately benefiting both the employees and the organization.

Working Code:

Library:

```
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt
from sklearn.preprocessing import LabelEncoder, StandardScaler
from sklearn.model_selection import train_test_split, GridSearchCV
from sklearn.metrics import accuracy_score, classification_report, confusion_matrix
import warnings
warnings.filterwarnings('ignore')
```

Import Data:

```
Work_data = pd.read_csv('C:\Users\hp\OneDrive\Desktop\Data.csv')
```

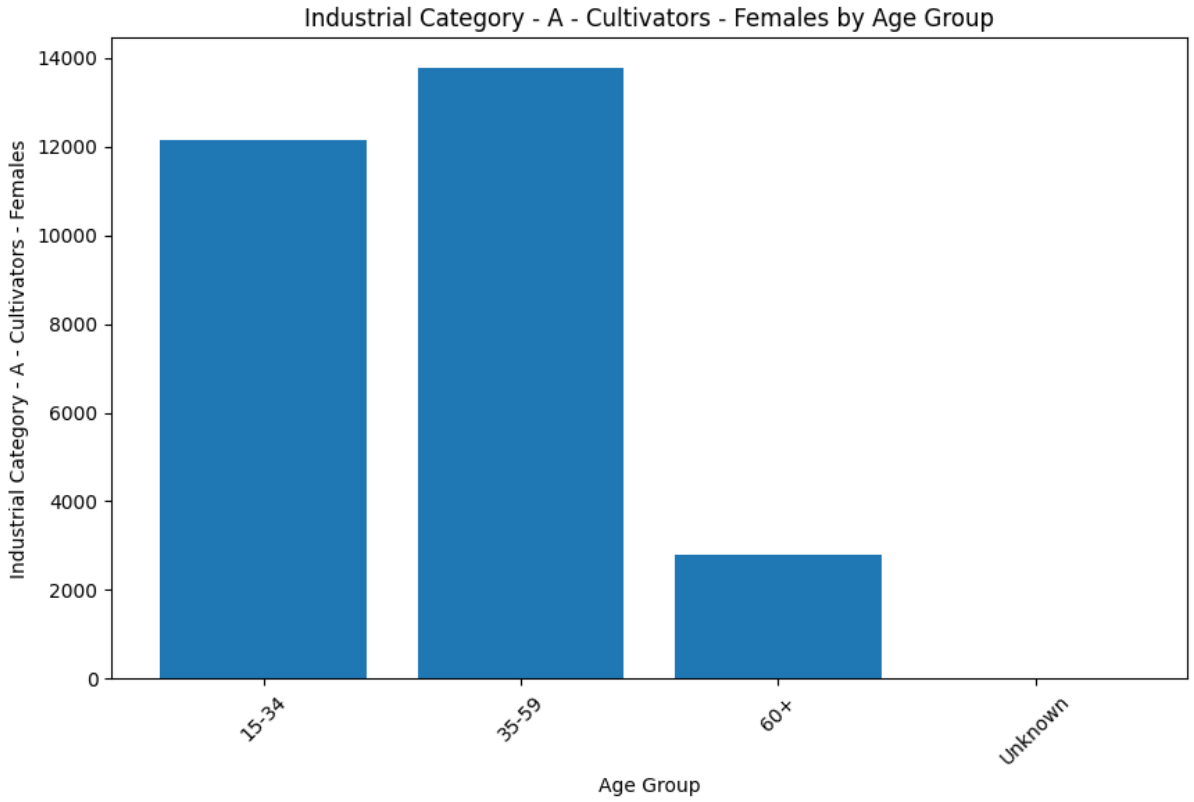
Analysis:

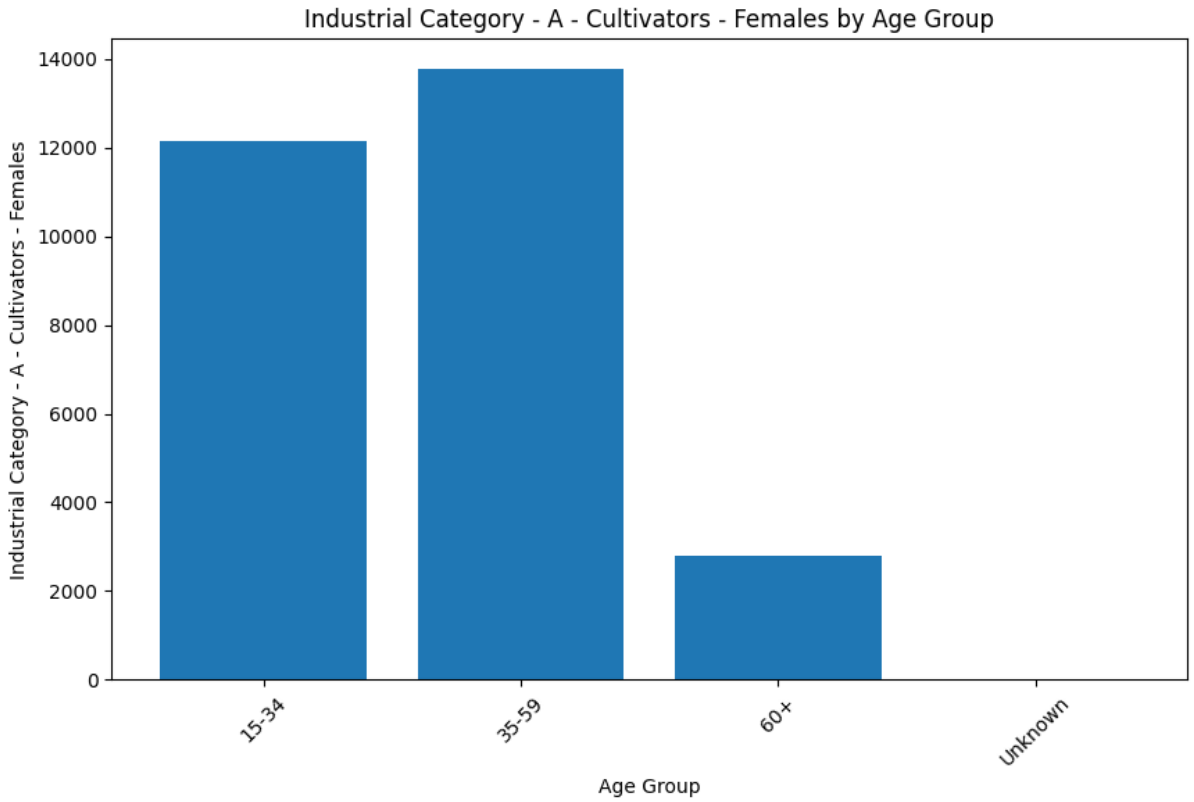
```
Data.shape
Data.columns
data.head()
data.info()
data.corr()
```

Graph Plot:

```
import pandas as pd
import matplotlib.pyplot as plt

# Load the preprocessed data
data = pd.read_csv("preprocessed_data.csv")

# List of columns to visualize
columns_to_visualize = [
    'Worked for 3 months or more but less than 6 months -  Persons',
    'Worked for 3 months or more but less than 6 months - Males',
    'Worked for 3 months or more but less than 6 months - Females',
    'Worked for less than 3 months - Persons',
    'Worked for less than 3 months - Males',
    'Worked for less than 3 months - Females',
    'Industrial Category - A - Cultivators - Persons',
    'Industrial Category - A - Cultivators - Males',
    'Industrial Category - A - Cultivators - Females',
```

```
    # Add the rest of the columns you want to visualize here
]

# Create a bar plot for each column
for column in columns_to_visualize:
    plt.figure(figsize=(10, 6))
    plt.bar(data['Age group'], data[column])
    plt.title(f'{column} by Age Group')
    plt.xlabel('Age Group')
    plt.ylabel(column)
    plt.xticks(rotation=45)
    plt.show()
```





Data Preprocessing Steps

Step 1: Data Loading

- The dataset was loaded from the provided CSV file.

```
import pandas as pd
data = pd.read_csv("data.csv")
```

INITIAL DATASET

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U
1	Table Code	State Code	District Cod	Area Name	Total/ Rural	Age group	Worked for	Worked for	Worked for	Worked for	Worked for	Worked for	Industrial Ca	Industrial Ca	Industrial Ca	Industrial Ca	Industrial Ca	Industrial Ca	Industrial Ca	Industrial Ca	Industri
2	B08065C	`33	`000	State - TAM	Total	Total	1200828	589003	611825	221386	99368	122018	64235	34632	29603	907752	404844	502908	29410	16268	13
3	B08065C	`33	`000	State - TAM	Total	`5-14	27791	14125	13666	2447	1247	1200	1710	825	885	6398	3130	3268	190	107	
4	B08065C	`33	`000	State - TAM	Total	15-34	514340	259560	254780	92423	43892	48531	24863	12711	12152	345420	152968	192452	9430	5443	3
5	B08065C	`33	`000	State - TAM	Total	35-59	542581	251957	290624	99202	40691	58511	29692	15927	13765	450052	192771	257281	15744	8230	7
6	B08065C	`33	`000	State - TAM	Total	60+	115103	62833	52270	27165	13465	13700	7930	5151	2779	105325	55730	49595	4028	2470	1
7	B08065C	`33	`000	State - TAM	Total	Age not stat	1013	528	485	149	73	76	40	18	22	557	245	312	18	18	
8	B08065C	`33	`000	State - TAM	Rural	Total	966645	459738	506907	174443	73663	100780	59637	32189	27448	824698	364131	460567	19758	11033	8
9	B08065C	`33	`000	State - TAM	Rural	`5-14	17239	8713	8526	1977	985	992	1443	684	759	6005	2922	3083	144	80	
10	B08065C	`33	`000	State - TAM	Rural	15-34	406847	198575	208272	71974	31917	40057	22933	11766	11167	316885	138622	178263	6687	3909	2
11	B08065C	`33	`000	State - TAM	Rural	35-59	444800	199573	245227	77922	29808	48114	27799	14887	12912	406147	172178	233969	10307	5468	4
12	B08065C	`33	`000	State - TAM	Rural	60+	97011	52498	44513	22446	10902	11544	7425	4835	2590	95151	50192	44959	2608	1564	1
13	B08065C	`33	`000	State - TAM	Rural	Age not stat	748	379	369	124	51	73	37	17	20	510	217	293	12	12	
14	B08065C	`33	`000	State - TAM	Urban	Total	234183	129265	104918	46943	25705	21238	4598	2443	2155	83054	40713	42341	9652	5235	4
15	B08065C	`33	`000	State - TAM	Urban	`5-14	10552	5412	5140	470	262	208	267	141	126	393	208	185	46	27	
16	B08065C	`33	`000	State - TAM	Urban	15-34	107493	60985	46508	20449	11975	8474	1930	945	985	28535	14346	14189	2743	1534	1
17	B08065C	`33	`000	State - TAM	Urban	35-59	97781	52384	45397	21280	10883	10397	1893	1040	853	43905	20593	23312	5437	2762	2

Step 2: Column Removal

- Unnecessary columns, namely "Table Code," "State Code," and "District Code," were removed to simplify the dataset and focus on relevant attributes.

```
columns_to_drop = ["Table Code", "State Code", "District Code"]
data.drop(columns=columns_to_drop, inplace=True)
```

OUTPUT:

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U
1	Area Name	Total/ Rural	Age group	Worked for	Worked for	Worked for	Worked for	Worked for	Worked for	Industrial Ca	Industrial Ca	Industrial Ca	Industrial Ca	Industrial Ca	Industrial Ca	Industrial Ca	Industrial Ca	Industrial Ca	Industrial Ca	Industrial Ca	Industr
2	State - TAM	Total	Total	1200828	589003	611825	221386	99368	122018	64235	34632	29603	907752	404844	502908	29410	16268	13142	2853	1862	
3	State - TAM	Total	'5-14	27791	14125	13666	2447	1247	1200	1710	825	885	6398	3130	3268	190	107	83	9	9	
4	State - TAM	Total	15-34	514340	259560	254780	92423	43892	48531	24863	12711	12152	345420	152968	192452	9430	5443	3987	1174	839	
5	State - TAM	Total	35-59	542581	251957	290624	99202	40691	58511	29692	15927	13765	450052	192771	257281	15744	8230	7514	1436	860	
6	State - TAM	Total	60+	115103	62833	52270	27165	13465	13700	7930	5151	2779	105325	55730	49595	4028	2470	1558	234	154	
7	State - TAM	Total	Age not stat	1013	528	485	149	73	76	40	18	22	557	245	312	18	18	0	0	0	
8	State - TAM	Rural	Total	966645	459738	506907	174443	73663	100780	59637	32189	27448	824698	364131	460567	19758	11033	8725	1728	1191	
9	State - TAM	Rural	'5-14	17239	8713	8526	1977	985	992	1443	684	759	6005	2922	3083	144	80	64	6	6	
10	State - TAM	Rural	15-34	406847	198575	208272	71974	31917	40057	22933	11766	11167	316885	138622	178263	6687	3909	2778	732	543	
11	State - TAM	Rural	35-59	444800	199573	245227	77922	29808	48114	27799	14887	12912	406147	172178	233969	10307	5468	4839	882	567	
12	State - TAM	Rural	60+	97011	52498	44513	22446	10902	11544	7425	4835	2590	95151	50192	44959	2608	1564	1044	108	75	
13	State - TAM	Rural	Age not stat	748	379	369	124	51	73	37	17	20	510	217	293	12	12	0	0	0	
14	State - TAM	Urban	Total	234183	129265	104918	46943	25705	21238	4598	2443	2155	83054	40713	42341	9652	5235	4417	1125	671	
15	State - TAM	Urban	'5-14	10552	5412	5140	470	262	208	267	141	126	393	208	185	46	27	19	3	3	
16	State - TAM	Urban	15-34	107493	60985	46508	20449	11975	8474	1930	945	985	28535	14346	14189	2743	1534	1209	442	296	
17	State - TAM	Urban	35-59	97781	52384	45397	21280	10883	10397	1893	1040	853	43905	20593	23312	5437	2762	2675	554	293	
18	State - TAM	Urban	60+	18092	10335	7757	4719	2563	2156	505	316	189	10174	5538	4636	1420	906	514	126	79	
19	State - TAM	Urban	Age not stat	265	149	116	25	22	3	3	1	2	47	28	19	6	6	0	0	0	
20	District - Thi	Total	Total	74448	39295	35153	15866	8004	7862	3066	1663	1403	42579	20345	22234	1519	1025	494	63	47	

Step 3: Removal of Rows with "Total" in Age Group

- Rows with "Total" in the "Age Group" column were eliminated as they represent aggregated data and do not pertain to individual age groups.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	
1	Area Name	Total/ Rural	Age group	Worked for	Worked for	Worked for	Worked for	Worked for	Worked for	Industrial Ca	Industrial Ca	Industrial Ca	Industrial Ca	Industrial Ca	Industrial Ca	Industrial Ca	Industrial Ca	Industrial Ca	Industrial Ca	Industrial Ca	Indus
2	State - TAM	Total	'5-14	27791	14125	13666	2447	1247	1200	1710	825	885	6398	3130	3268	190	107	83	9	9	
3	State - TAM	Total	15-34	514340	259560	254780	92423	43892	48531	24863	12711	12152	345420	152968	192452	9430	5443	3987	1174	839	
4	State - TAM	Total	35-59	542581	251957	290624	99202	40691	58511	29692	15927	13765	450052	192771	257281	15744	8230	7514	1436	860	
5	State - TAM	Total	60+	115103	62833	52270	27165	13465	13700	7930	5151	2779	105325	55730	49595	4028	2470	1558	234	154	
6	State - TAM	Total	Age not stat	1013	528	485	149	73	76	40	18	22	557	245	312	18	18	0	0	0	
7	State - TAM	Rural	'5-14	17239	8713	8526	1977	985	992	1443	684	759	6005	2922	3083	144	80	64	6	6	
8	State - TAM	Rural	15-34	406847	198575	208272	71974	31917	40057	22933	11766	11167	316885	138622	178263	6687	3909	2778	732	543	
9	State - TAM	Rural	35-59	444800	199573	245227	77922	29808	48114	27799	14887	12912	406147	172178	233969	10307	5468	4839	882	567	
10	State - TAM	Rural	60+	97011	52498	44513	22446	10902	11544	7425	4835	2590	95151	50192	44959	2608	1564	1044	108	75	
11	State - TAM	Rural	Age not stat	748	379	369	124	51	73	37	17	20	510	217	293	12	12	0	0	0	
12	State - TAM	Urban	'5-14	10552	5412	5140	470	262	208	267	141	126	393	208	185	46	27	19	3	3	
13	State - TAM	Urban	15-34	107493	60985	46508	20449	11975	8474	1930	945	985	28535	14346	14189	2743	1534	1209	442	296	
14	State - TAM	Urban	35-59	97781	52384	45397	21280	10883	10397	1893	1040	853	43905	20593	23312	5437	2762	2675	554	293	
15	State - TAM	Urban	60+	18092	10335	7757	4719	2563	2156	505	316	189	10174	5538	4636	1420	906	514	126	79	
16	State - TAM	Urban	Age not stat	265	149	116	25	22	3	3	1	2	47	28	19	6	6	0	0	0	
17	District - Thi	Total	'5-14	2521	1284	1237	147	82	65	122	56	66	330	154	176	12	12	0	0	0	
18	District - Thi	Total	15-34	33568	18049	15519	6529	3654	2875	1225	632	593	15591	7257	8334	570	387	183	27	21	
19	District - Thi	Total	35-59	32568	16771	15797	7718	3529	4189	1414	792	622	22192	10446	11746	788	532	256	36	26	
20	District - Thi	Total	60+	5716	3147	2569	1465	739	726	305	183	122	4441	2476	1965	149	94	55	0	0	

Step 4: Age Group Mapping

- The "Age Group" column was standardized to ensure a consistent format.
- Specific mappings:
  - "5-14" remains "5-14"
  - "15-34" remains "15-34"
  - "35-59" remains "35-59"

- "60+" remains "60+"
- "Age not stated" has been replaced with "Unknown" to address missing or unspecified age group entries.

```
age_group_mapping = {
    '5-14': '5-14',
    '15-34': '15-34',
    '35-59': '35-59',
    '60+': '60+',
    'Age not stated': 'Unknown',

}
data['Age group'] = data['Age group'].map(age_group_mapping)
data['Age group'] = data['Age group'].astype('category')
```

OUTPUT:

Area Name	Total/ Rural	Age group	Worked for	Worked for	Worked for	Worked for	Worked for	Worked for	Industrial Ca	Industrial Ca	Industrial Ca	Industrial Ca	Industrial Ca	Industrial Ca	Industrial Ca	Industrial Ca	Industrial Ca	Industrial Ca	Industri
State - TAM	Total		27791	14125	13666	2447	1247	1200	1710	825	885	6398	3130	3268	190	107	83	9	9
State - TAM	Total	15-34	514340	259560	254780	92423	43892	48531	24863	12711	12152	345420	152968	192452	9430	5443	3987	1174	839
State - TAM	Total	35-59	542581	251957	290624	99202	40691	58511	29692	15927	13765	450052	192771	257281	15744	8230	7514	1436	860
State - TAM	Total	60+	115103	62833	52270	27165	13465	13700	7930	5151	2779	105325	55730	49595	4028	2470	1558	234	154
State - TAM	Total	Unknown	1013	528	485	149	73	76	40	18	22	557	245	312	18	18	0	0	0
State - TAM	Rural		17239	8713	8526	1977	985	992	1443	684	759	6005	2922	3083	144	80	64	6	6
State - TAM	Rural	15-34	406847	198575	208272	71974	31917	40057	22933	11766	11167	316885	138622	178263	6687	3909	2778	732	543
State - TAM	Rural	35-59	444800	199573	245227	77922	29808	48114	27799	14887	12912	406147	172178	233969	10307	5468	4839	882	567
State - TAM	Rural	60+	97011	52498	44513	22446	10902	11544	7425	4835	2590	95151	50192	44959	2608	1564	1044	108	75
State - TAM	Rural	Unknown	748	379	369	124	51	73	37	17	20	510	217	293	12	12	0	0	0
State - TAM	Urban		10552	5412	5140	470	262	208	267	141	126	393	208	185	46	27	19	3	3
State - TAM	Urban	15-34	107493	60985	46508	20449	11975	8474	1930	945	985	28535	14346	14189	2743	1534	1209	442	296
State - TAM	Urban	35-59	97781	52384	45397	21280	10883	10397	1893	1040	853	43905	20593	23312	5437	2762	2675	554	293
State - TAM	Urban	60+	18092	10335	7757	4719	2563	2156	505	316	189	10174	5538	4636	1420	906	514	126	79
State - TAM	Urban	Unknown	265	149	116	25	22	3	3	1	2	47	28	19	6	6	0	0	0
District - Thi	Total		2521	1284	1237	147	82	65	122	56	66	330	154	176	12	12	0	0	0
District - Thi	Total	15-34	33568	18049	15519	6529	3654	2875	1225	632	593	15591	7257	8334	570	387	183	27	21
District - Thi	Total	35-59	32568	16771	15797	7718	3529	4189	1414	792	622	22192	10446	11746	788	532	256	36	26
District - Thi	Total	60+	5716	3147	2569	1465	739	726	305	183	122	4441	2476	1965	149	94	55	0	0

Step 5: Removal of Rows with Empty Entries in Age Group and Data Saving

- Rows containing empty or missing entries in the "Age Group" column were removed to maintain data consistency.
- The preprocessed data was saved to a new CSV file, "preprocessed_data.csv," making it suitable for further analysis.

```
data = data.dropna(subset=['Age group'])
data.to_csv("preprocessed_data.csv", index=False)
```

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U
1	Area Name	Total/ Rural	Age group	Worked for	Worked for	Worked for	Worked for	Worked for	Worked for	Industrial Ca	Industrial Ca	Industrial Ca	Industrial Ca	Industrial Ca	Industrial Ca	Industrial Ca	Industrial Ca	Industrial Ca	Industrial Ca	Industrial Ca	Industri
2	State - TAM	Total	15-34	514340	259560	254780	92423	43892	48531	24863	12711	12152	345420	152968	192452	9430	5443	3987	1174	839	
3	State - TAM	Total	35-59	542581	251957	290624	99202	40691	58511	29692	15927	13765	450052	192771	257281	15744	8230	7514	1436	860	
4	State - TAM	Total	60+	115103	62833	52270	27165	13465	13700	7930	5151	2779	105325	55730	49595	4028	2470	1558	234	154	
5	State - TAM	Total	Unknown	1013	528	485	149	73	76	40	18	22	557	245	312	18	18	0	0	0	
6	State - TAM	Rural	15-34	406847	198575	208272	71974	31917	40057	22933	11766	11167	316885	138622	178263	6687	3909	2778	732	543	
7	State - TAM	Rural	35-59	444800	199573	245227	77922	29808	48114	27799	14887	12912	406147	172178	233969	10307	5468	4839	882	567	
8	State - TAM	Rural	60+	97011	52498	44513	22446	10902	11544	7425	4835	2590	95151	50192	44959	2608	1564	1044	108	75	
9	State - TAM	Rural	Unknown	748	379	369	124	51	73	37	17	20	510	217	293	12	12	0	0	0	
10	State - TAM	Urban	15-34	107493	60985	46508	20449	11975	8474	1930	945	985	28535	14346	14189	2743	1534	1209	442	296	
11	State - TAM	Urban	35-59	97781	52384	45397	21280	10883	10397	1893	1040	853	43905	20593	23312	5437	2762	2675	554	293	
12	State - TAM	Urban	60+	18092	10335	7757	4719	2563	2156	505	316	189	10174	5538	4636	1420	906	514	126	79	
13	State - TAM	Urban	Unknown	265	149	116	25	22	3	3	1	2	47	28	19	6	6	0	0	0	
14	District - Thi	Total	15-34	33568	18049	15519	6529	3654	2875	1225	632	593	15591	7257	8334	570	387	183	27	21	
15	District - Thi	Total	35-59	32568	16771	15797	7718	3529	4189	1414	792	622	22192	10446	11746	788	532	256	36	26	
16	District - Thi	Total	60+	5716	3147	2569	1465	739	726	305	183	122	4441	2476	1965	149	94	55	0	0	
17	District - Thi	Total	Unknown	75	44	31	7	0	7	0	0	0	25	12	13	0	0	0	0	0	
18	District - Thi	Rural	15-34	23965	12377	11588	4713	2443	2270	1109	566	543	14527	6700	7827	484	326	158	18	12	
19	District - Thi	Rural	35-59	25421	12417	13004	6104	2574	3530	1320	740	580	20723	9675	11048	650	444	206	20	14	
20	District - Thi	Rural	60+	4718	2516	2202	1193	575	618	276	162	114	4175	2327	1848	116	70	46	0	0	

Code Explanation

The project code begins by loading the dataset and performing data preprocessing. Irrelevant columns are removed, the 'Total' category is filtered out from the 'Age group' column, age group categories are mapped to descriptive labels, and missing values are handled to ensure data quality and accuracy.

The subsequent data visualization phase is critical for the project's success. The chosen visualization types help convey complex information in a clear and accessible manner. These visualizations, including bar plots, histograms, box plots, scatter plots, and a heatmap, offer unique insights into the demographic characteristics of marginal workers.

Project Scope

The scope of this project encompasses several key aspects:

1. Data Collection: Acquiring and accessing the relevant dataset containing information about marginal workers in Tamil Nadu.

2. Data Preprocessing: Preparing the dataset for analysis by cleaning and structuring the data. This involves removing irrelevant columns and handling missing or erroneous data points. Additionally, filtering out the 'Total' category from the 'Age group' column to focus on specific age groups.

3. Data Analysis: Delving into the dataset to extract insights about the demographic characteristics of marginal workers. This phase involves statistical calculations and exploration of patterns and trends.

4. Data Visualization: Creating visualizations that vividly represent the findings of the analysis. These visualizations are a critical component of the project, as they make complex data more understandable and help stakeholders grasp the key insights.

Approach
The project follows a systematic approach:

1. Data Collection:
- The dataset is obtained from [Specify the data source]. This dataset contains information about marginal workers in Tamil Nadu, covering a range of demographic variables and categories.

2. Data Preprocessing:
- Irrelevant columns such as 'Table Code,' 'State Code,' and 'District Code' are removed to focus on essential attributes.
- 'Total' rows in the 'Age group' column are filtered out, as the analysis is intended to be age-specific.
- Age group categories are mapped to more descriptive labels to improve clarity and understanding.
- Missing values are handled to ensure data integrity.

Data Visualizations

1. Bar Plots:
- Bar plots are generated to visualize the distribution of various characteristics of marginal workers segmented by age groups. These plots offer a clear and concise representation of data distribution.

2. Histogram:
- A histogram is created to illustrate the distribution of individuals who worked for 3-6 months. This histogram provides an overview of the distribution of this specific category.

3. Box Plot:
- A box plot is generated to show the distribution of individuals who worked for 3-6 months, segmented by age groups. It helps identify variations in different age groups.

4. Scatter Plot:
- A scatter plot is produced to reveal the relationship between the number of individuals who worked for 3-6 months and the number of males in that category. Scatter plots are useful for understanding correlations.

5. Heatmap:

- A heatmap is designed to visualize correlations between selected columns. This heatmap provides insights into the relationships between key variables.

Conclusion:

In conclusion, the study of the demographics and challenges faced by marginalized workers in Tamil Nadu has illuminated the pressing need for innovative assessment methodologies and solutions. Leveraging a data-driven approach and design thinking principles, this research not only deepened our understanding of the multifaceted issues confronting these workers but also highlighted the pivotal role of innovation in addressing their complex socio-economic and educational hurdles. By applying technological solutions, such as advanced data analytics, skill development platforms, and blockchain for labor rights, we can accurately assess the needs of marginalized workers, paving the way for more empathetic and effective solutions to enhance their well-being and promote social inclusion in Tamil Nadu. This comprehensive approach underscores the importance of a holistic and empathetic understanding of the challenges faced by marginalized workers in the region, emphasizing the potential for positive change through innovative solutions.

Team Members:

Priyadarshini(2021115078)
Rakshitha Senthil(2021115082)
Ramkrishna(2021115083)
Ranjith Kumar(2021115084)
Saiganesh(2021115316)