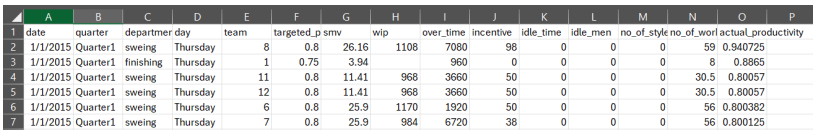
Data Collection and Preprocessing Phase

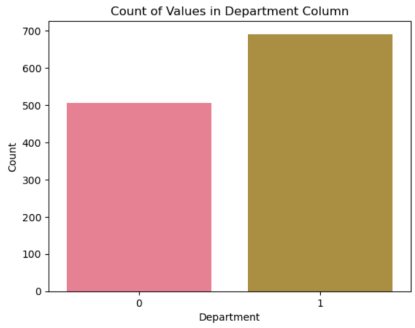
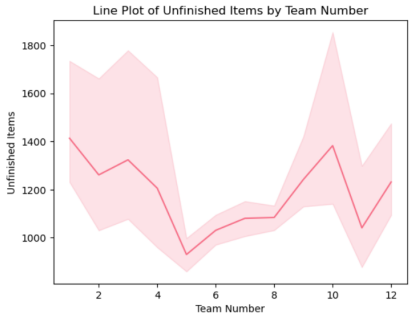
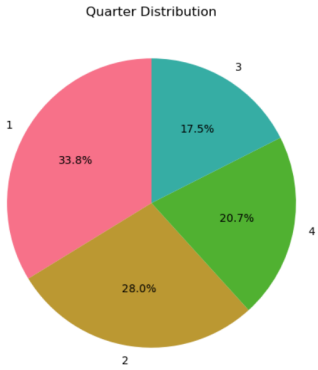
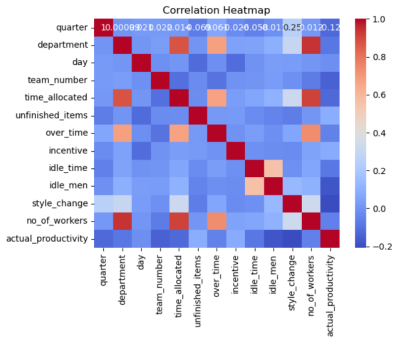
Date	15 March 2024
Team ID	SWTID1720673861
Project Title	Garment Worker efficiency Calculator
Maximum Marks	6 Marks

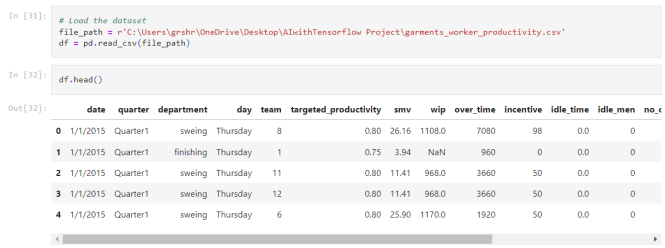
Data Exploration and Preprocessing Report

Dataset variables will be statistically analyzed to identify patterns and outliers, with Python employed for preprocessing tasks like normalization and feature engineering. Data cleaning will address missing values and outliers, ensuring quality for subsequent analysis and modeling, and forming a strong foundation for insights and predictions.

Section	Description
Data Overview	Dimension: 1197 rows × 15 columns Descriptive statistics: 
Univariate Analysis	

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	date	quarter	departmen	day	team	targeted_p	smv	wip	over_time	incentive	idle_time	idle_men	no_of_style	no_of_worl	actual_productivity	
2	1/1/2015	Quarter1	sweing	Thursday	8	0.8	26.16	1108	7080	98	0	0	0	59	0.940725	
3	1/1/2015	Quarter1	finishing	Thursday	1	0.75	3.94		960	0	0	0	0	8	0.8865	
4	1/1/2015	Quarter1	sweing	Thursday	11	0.8	11.41	968	3660	50	0	0	0	30.5	0.80057	
5	1/1/2015	Quarter1	sweing	Thursday	12	0.8	11.41	968	3660	50	0	0	0	30.5	0.80057	
6	1/1/2015	Quarter1	sweing	Thursday	6	0.8	25.9	1170	1920	50	0	0	0	56	0.800382	
7	1/1/2015	Quarter1	sweing	Thursday	7	0.8	25.9	984	6720	38	0	0	0	56	0.800125	

	
Bivariate Analysis	
Multivariate Analysis	 

Outliers and Anomalies	-
Data Preprocessing Code Screenshots	
Loading Data	 date quarter department day team targeted_productivity smv wip over_time incentive idle_time idle_men no_c 0 1/1/2015 Quarter1 sweing Thursday 8 0.80 26.16 1108.0 7080 98 0.0 0 1 1/1/2015 Quarter1 finishing Thursday 1 0.75 3.94 NaN 960 0 0.0 0 2 1/1/2015 Quarter1 sweing Thursday 11 0.80 11.41 968.0 3660 50 0.0 0 3 1/1/2015 Quarter1 sweing Thursday 12 0.80 11.41 968.0 3660 50 0.0 0 4 1/1/2015 Quarter1 sweing Thursday 6 0.80 25.90 1170.0 1920 50 0.0 0
Handling Missing Data	
Data Transformation	
Feature Engineering	Attached the codes in final submission.
Save Processed Data	-