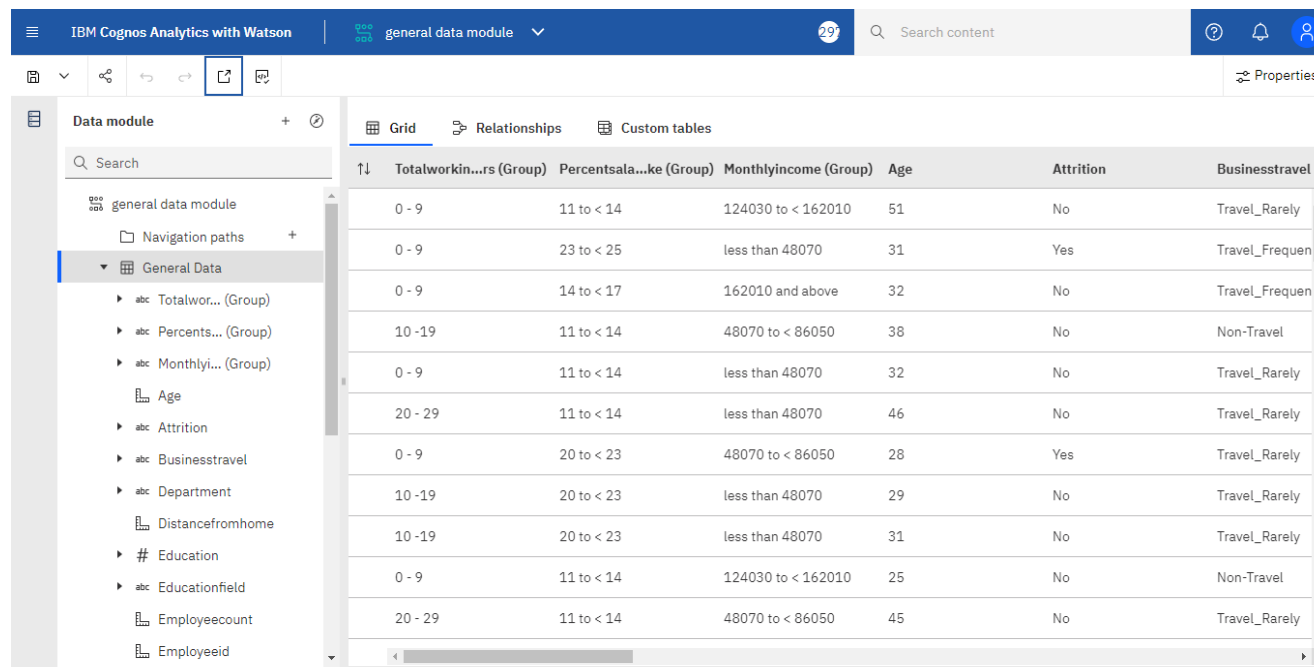


Final Deliverables

Data module

Date	19 November 2022
Team ID	PNT2022TMID40488
Project Name	Corporate Employee Attrition Analysis

GENERAL DATA:



The screenshot displays the IBM Cognos Analytics interface for the 'general data module'. The left sidebar shows a tree view with 'General Data' expanded, listing various data sources. The main area shows a table with columns: Totalworkin...rs (Group), Percentsala...ke (Group), Monthlyincome (Group), Age, Attrition, and Businesstravel. The table contains 10 rows of data.

Totalworkin...rs (Group)	Percentsala...ke (Group)	Monthlyincome (Group)	Age	Attrition	Businesstravel
0 - 9	11 to < 14	124030 to < 162010	51	No	Travel_Rarely
0 - 9	23 to < 25	less than 48070	31	Yes	Travel_Frequen
0 - 9	14 to < 17	162010 and above	32	No	Travel_Frequen
10 -19	11 to < 14	48070 to < 86050	38	No	Non-Travel
0 - 9	11 to < 14	less than 48070	32	No	Travel_Rarely
20 - 29	11 to < 14	less than 48070	46	No	Travel_Rarely
0 - 9	20 to < 23	48070 to < 86050	28	Yes	Travel_Rarely
10 -19	20 to < 23	less than 48070	29	No	Travel_Rarely
10 -19	20 to < 23	less than 48070	31	No	Travel_Rarely
0 - 9	11 to < 14	124030 to < 162010	25	No	Non-Travel
20 - 29	11 to < 14	48070 to < 86050	45	No	Travel_Rarely

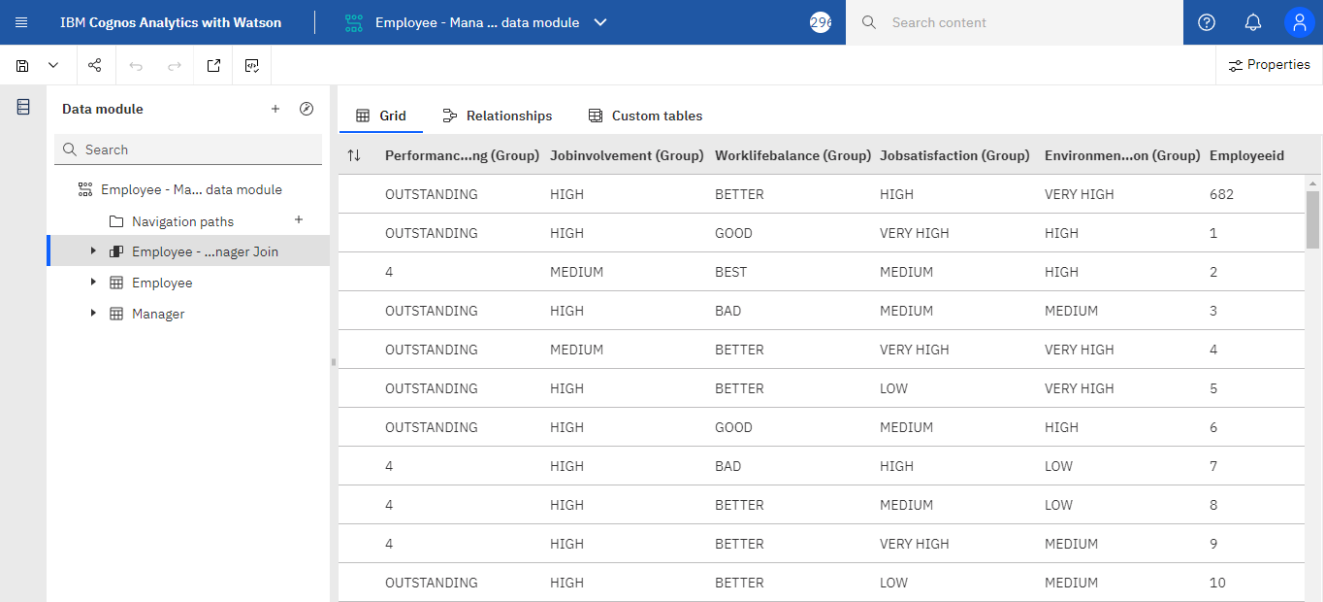
The given dataset is general_data.csv:

- The given data is uploaded from the IBM DB2 to IBM Cognos.
- The data that are uploaded in schema is loaded from the IBM DB2 and they are processed.
- The general_data.csv file that is loaded is a raw data file.
- The data module is created from the raw data file.
- The null values are replaced.
- The columns having continuous range of values are grouped to form another set of grouped column.
- Then the cleaned module is created.

Link:

https://us3.ca.analytics.ibm.com/bi/?perspective=ca-modeller&pathRef=.my_folders%2FFinal%2Fgeneral%2Bdata%2Bmodule

EMPLOYEE – MANAGER DATA:



Performance...ng (Group)	Jobinvolvement (Group)	Worklifebalance (Group)	Jobsatisfaction (Group)	Environmen...on (Group)	Employeeid
OUTSTANDING	HIGH	BETTER	HIGH	VERY HIGH	682
OUTSTANDING	HIGH	GOOD	VERY HIGH	HIGH	1
4	MEDIUM	BEST	MEDIUM	HIGH	2
OUTSTANDING	HIGH	BAD	MEDIUM	MEDIUM	3
OUTSTANDING	MEDIUM	BETTER	VERY HIGH	VERY HIGH	4
OUTSTANDING	HIGH	BETTER	LOW	VERY HIGH	5
OUTSTANDING	HIGH	GOOD	MEDIUM	HIGH	6
4	HIGH	BAD	HIGH	LOW	7
4	HIGH	BETTER	MEDIUM	LOW	8
4	HIGH	BETTER	VERY HIGH	MEDIUM	9
OUTSTANDING	HIGH	BETTER	LOW	MEDIUM	10

The given dataset is employee_survey_data.csv & manager_survey_data.csv:

- The given data is uploaded from the IBM DB2 to IBM Cognos.
- The data that are uploaded in schema is loaded from the IBM DB2 and they are processed.
- The employee_survey_data.csv & manager_survey_data.csv file that is loaded is a raw data file.
- Both the csv files has the same column in the datasets.
- Thus the csv files are merged into one single dataset by the join condition.
- The join condition is achieved in raw data file.
- The data module is created from the raw data file.
- The null values are replaced.
- The columns having continuous range of values are grouped to form another set of grouped column.
- Then the cleaned module is created.

Link:

https://us3.ca.analytics.ibm.com/bi/?perspective=ca-modeller&pathRef=.my_folders%2FFinal%2FEmployee%2B-%2BManager%2Bdata%2Bmodule