

Black Women Data 2022

Power BI Workshop

Case: You are a consultant working with a human capital division that provides people analytics services to Fortune 500 organizations. The organization has asked you to build a dashboard around their team and organizational performance and employee sentiment. This case is inspired by this [Kaggle challenge](#).

Data: [Introduction to the HR Dataset - Version 14](#) - Last Updated April, 2021. This HR Dataset is a synthetic data set created specifically to use for HR analytics cases and is updated every year or so, by the owners. Updates include additional columns, and to make slight changes to the underlying data.


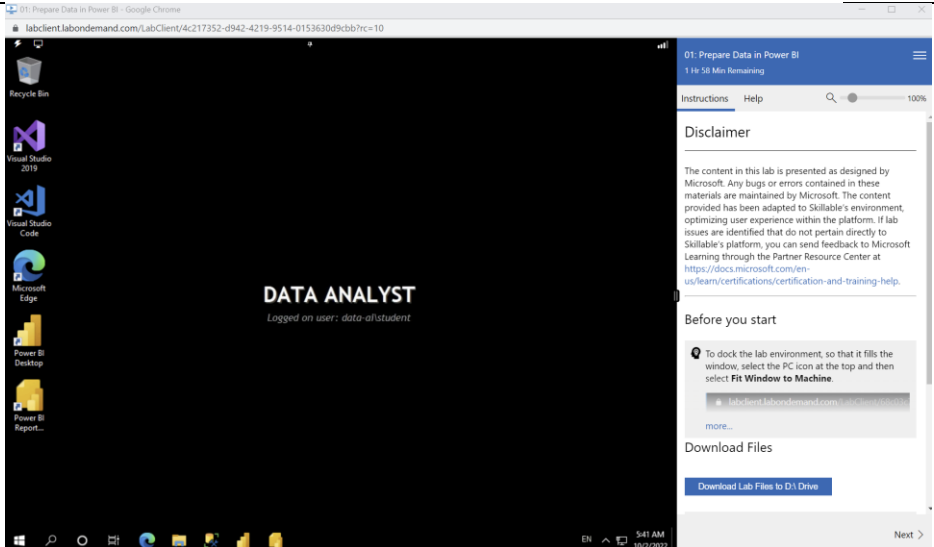
Inspirational Questions: Here are some open-ended questions that you can explore and try to address through creating visualizations, or R or Python analyses.

- *Is there any relationship between who a person works for and their performance score?*
- *What is the overall diversity profile of the organization?*
- *What are our best recruiting sources if we want to ensure a diverse organization?*
- *Can we predict who is going to terminate and who isn't? What level of accuracy can we achieve on this?*
- *Are there areas of the company where pay is not equitable?*
- Is there any relationship between who a person works for and their performance score?
- What is the overall diversity profile of the organization?
- What are our best recruiting sources if we want to ensure a diverse organization?
- Can we predict who is going to terminate and who isn't? What level of accuracy can we achieve on this?
- Are there areas of the company where pay is not equitable?
- Is there a relationship between age and performance
- Does working on special projects affects performance

Data Dictionary:

Feature	Description	DataType
Employee Name	Employee's full name	Text
EmpID	Employee ID is unique to each employee	Text
MarriedID	Is the person married (1 or 0 for yes or no)	Binary
MaritalStatusID	Marital status code that matches the text field MaritalDesc	Integer
EmpStatusID	Employment status code that matches text field EmploymentStatus	Integer
DeptID	Department ID code that matches the department the employee works in	Integer
PerfScoreID	Performance Score code that matches the employee's most recent performance score	Integer
FromDiversityJobFairID	Was the employee sourced from the Diversity job fair? 1 or 0 for yes or no	Binary
Salary	The person's yearly salary. \$ U.S. Dollars	Float
Termd	Has this employee been terminated - 1 or 0	Binary
PositionID	An integer indicating the person's position	Integer
Position	The text name/title of the position the person has	Text
State	The state that the person lives in	Text
Zip	The zip code for the employee	Text
DOB	Date of Birth for the employee	Date
Sex	Sex - M or F	Text
MaritalDesc	The marital status of the person (divorced, single, widowed, separated, etc)	Text
CitizenDesc	Label for whether the person is a Citizen or Eligible NonCitizen	Text
HispanicLatino	Yes or No field for whether the employee is Hispanic/Latino	Text
RaceDesc	Description/text of the race the person identifies with	Text
DateofHire	Date the person was hired	Date
DateofTermination	Date the person was terminated, only populated if, in fact, Termd = 1	Date
TermReason	A text reason / description for why the person was terminated	Text
EmploymentStatus	A description/category of the person's employment status. Anyone currently working full time = Active	Text
Department	Name of the department that the person works in	Text
ManagerName	The name of the person's immediate manager	Text
ManagerID	A unique identifier for each manager.	Integer
RecruitmentSource	The name of the recruitment source where the employee was recruited from	Text
PerformanceScore	Performance Score text/category (Fully Meets, Partially Meets, PIP, Exceeds)	Text
EngagementSurvey	Results from the last engagement survey, managed by our external partner	Float
EmpSatisfaction	A basic satisfaction score between 1 and 5, as reported on a recent employee satisfaction survey	Integer
SpecialProjectsCount	The number of special projects that the employee worked on during the last 6 months	Integer
LastPerformanceReviewDate	The most recent date of the person's last performance review.	Date
DaysLateLast30	The number of times that the employee was late to work during the last 30 days	Integer
Absences	The number of times the employee was absent from work.	Integer

Power BI Steps:

<p>Download all files:</p> <p>Files are available on Github</p> <p>Unzip files to Desktop</p>	
<p>Opening Power BI Options</p> <p>Option 1:</p> <ul style="list-style-type: none"> • Open Microsoft Virtual Machine • Visit Become a Power BI Analyst Path • Open up this lab: <ul style="list-style-type: none"> ○ https://learn.microsoft.com/en-us/training/modules/get-data/lab-prepare • Sign in using your own account or a training account • Click on Start Lab • Lab environment will begin to build and open up to a Windows environment where Power BI is readily available 	<div data-bbox="589 772 1409 1087"> <h3>Lab - Prepare data in Power BI Desktop</h3> <p>45 minutes 100 XP</p> <p>This unit includes a lab to complete.</p> <p>Use the free resources provided in the lab to complete the exercises in this unit. You will not be charged.</p> <p>Microsoft provides this lab experience and related content for educational purposes. All presented information is owned by Microsoft and intended solely for learning about the covered products and services in this Microsoft Learn module.</p> <p>Sign in to launch the lab</p> </div> <div data-bbox="589 1123 1008 1346"> <h3>Access your environment</h3> <p> Your lab environment is being built</p> <p>Your lab will be ready in about 2 minutes.</p> <div></div> </div>
<p>Lab Environment will look like this. The environment expires in 4 hours. If you would like to save your work, I suggest you log into a browser and sign into your Google drive and save your Power BI package there.</p>	

Option 2:

- Download Power BI
- <https://www.microsoft.com/en-us/download/details.aspx?id=58494>



Microsoft Power BI Desktop

With the Power BI Desktop you can visually explore your data through a free-form drag-and-drop canvas, a broad range of modern data visualizations, and an easy-to-use report authoring experience.

Download >

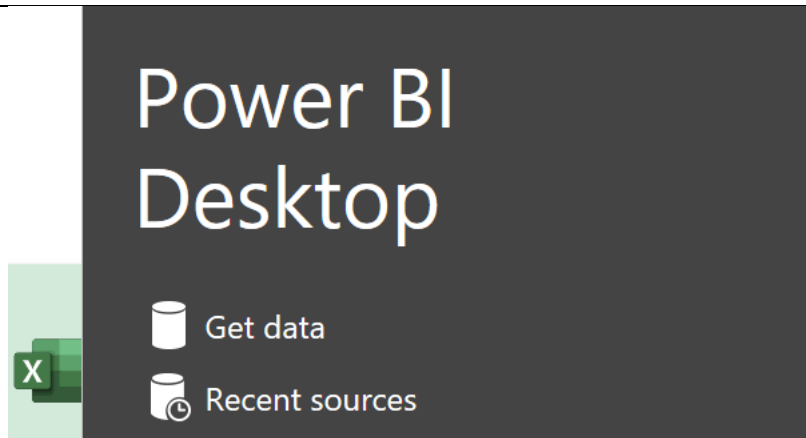
Advanced download options >

If you do not have a Microsoft account and want to use the lab without creating one, here are training accounts you can access.


<https://docs.google.com/spreadsheets/d/15ckDqvglrplwikV-NkpA5QJ5Eb876Aa0HUK7CV-jEE/edit?usp=sharing>


Open Power BI and Get Data

- To simplify this process we will use one data source: HR Dataset v14
- Select Text/CSV
- In the File Dialog, navigate to **HRDataset_v14**
 - You should only use the Excel Workbook option should you fall behind and need the tables we will build in the lab



Name

 HRData_Employees

 HRDataset_v14

Preview the file then click Transform Data to begin data preparation

HRDataset_v14.csv

File Origin: 65001: Unicode (UTF-8) | Delimiter: Comma | Data Type Detection: Based on first 200 rows

Employee_Name	EmpID	MarriedID	MaritalStatusID	GenderID	EmpStatusID	DeptID	PerfScoreID	FromDiversityJobFairID
Adinolfi, Wilson K	10026	0	0	1	1	5	4	0
Ait Sidi, Karthikeyan	10084	1	1	1	5	3	3	0
Akinkuolie, Sarah	10196	1	1	0	5	5	3	0
Alagbe, Trina	10088	1	1	0	1	5	3	0
Anderson, Carol	10069	0	2	0	5	5	3	0
Anderson, Linda	10002	0	0	0	1	5	4	0
Andreola, Colby	10194	0	0	0	1	4	3	0
Athwal, Sam	10062	0	4	1	1	5	3	0
Bachiochi, Linda	10114	0	0	0	3	5	3	1
Bacong, Alejandro	10250	0	2	1	1	3	3	0
Baczinski, Rachael	10252	1	1	0	5	5	3	1
Barbara, Thomas	10242	1	1	1	5	5	3	1
Barbossa, Hector	10012	0	2	1	1	3	4	1
Barone, Francesco A	10265	0	0	1	1	5	3	0
Barton, Nader	10066	0	2	1	5	5	3	0
Bates, Norman	10061	0	0	1	4	5	3	0
Beak, Kimberly	10023	1	1	0	2	5	4	0
Beatrice, Courtney	10055	0	0	0	1	5	3	0
Becker, Renee	10245	0	0	0	4	3	3	0
Becker, Scott	10277	0	0	1	3	5	3	0

Extract Table Using Examples | Load | Transform Data | Cancel

Begin with duplicating the table. Right click on the table and click **Duplicate Table**

Queries [1]

HRDataset_v14

- Copy
- Paste
- Delete
- Rename
- ☒ Enable load
- ☒ Include in report refresh
- Duplicate
- Reference
- Move To Group
- Move Up
- Move Down
- Create Function...
- Convert To Parameter
- Advanced Editor
- Properties...

16 Bates, Norman

Right Click on the table and name this Manager

Remove other columns except ManagerName, Manager ID and EmpID

name", type text}, {"EmpID", Int64.Type}, {"MarriedID", Int64.T

ManagerName	ManagerID	EmpID	MarriedID
Michael Albert			
Simon Roup			
Kissy Sullivan			
Elijah Gray			
Webster Butler			
Amy Dunn			
Alex Sweetwater			
Ketsia Liebig			
Brannon Miller			
Peter Monroe			
David Stanley			
Kissy Sullivan			
Simon Roup			
Kelley Spirea			
Michael Albert			
Kelley Spirea		18	Google Sear
Kelley Spirea		18	Employee R
Elijah Gray		16	Google Sear

Go to the original table and rename this table from HR Dataset v14 to Employees

Queries [7]

Employees

Remove ManagerName from Employees

ManagerName
Michael Albert
Simon Roup
Kissy Sullivan

123	PositionID	ABC	Position
0	19		Production Technician
1	27		Sr. DBA
1	20		Production Technician
0	19		Production Technician
1	19		Production Technician
0	19		Production Technician
0	24		Software Engineer
0	19		Production Technician
0	19		Production Technician
0	14		IT Support
1	19		Production Technician
1	19		Production Technician
0	9		Data Analyst
0	19		Production Technician
1	19		Production Technician
1	19		Production Technician I

Delete diversity data fields from original data table

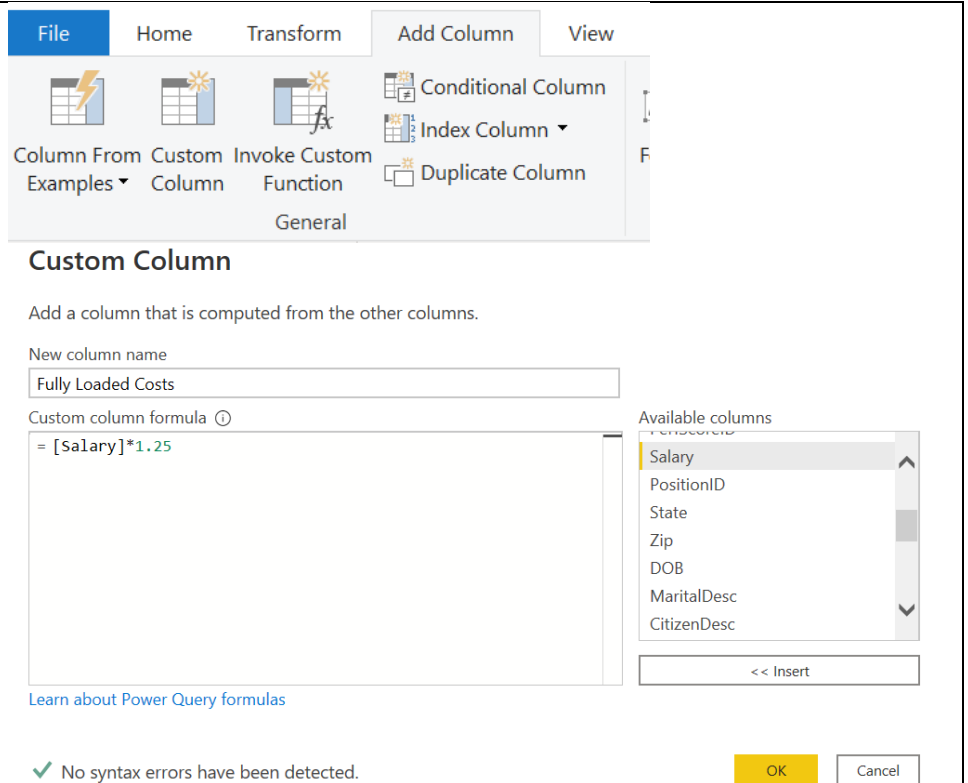
The screenshot shows a context menu in Power BI Desktop. The menu is open over a column named 'HispanicLat' in a table. The menu options are:

- Copy
- Remove Columns
- Remove Other Columns
- Add Column From Examples...
- Remove Duplicates
- Remove Errors
- Replace Values...
- Fill
- Change Type
- Merge Columns
- Group By...
- Unpivot Columns
- Unpivot Other Columns
- Unpivot Only Selected Columns
- Move

The background shows a table with columns 'HispanicLat' and 'White'. The 'HispanicLat' column contains values 'No' and 'Yes'. The 'White' column contains values 'White'.

Go back to Employees table and create a custom field.

Add Fully Loaded Cost



Custom Column

Add a column that is computed from the other columns.

New column name
Fully Loaded Costs

Custom column formula ①
= [Salary]*1.25

Available columns
Salary
PositionID
State
Zip
DOB
MaritalDesc
CitizenDesc

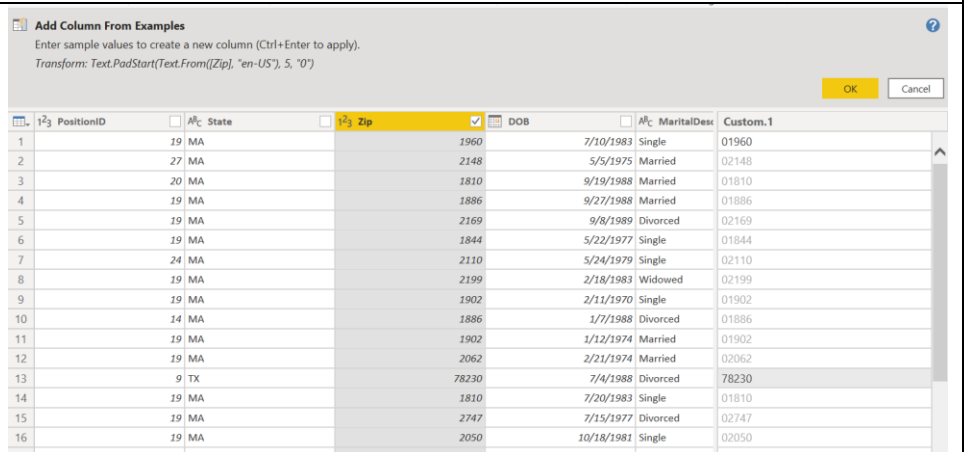
<< Insert

✓ No syntax errors have been detected.

OK Cancel

Update Zip Code column to ensure leading zero is back. We will use the Column from Examples feature.

Name this column Zip_c



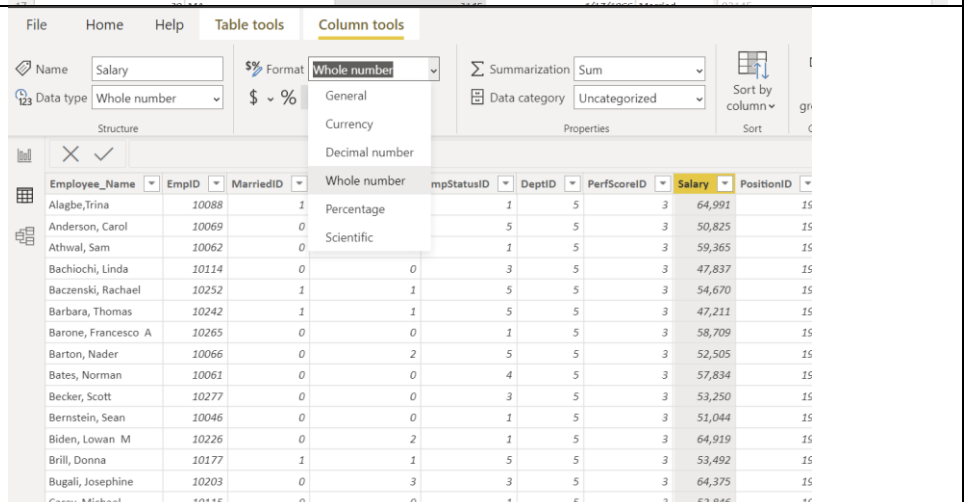
Add Column From Examples

Enter sample values to create a new column (Ctrl+Enter to apply).
Transform: Text.PadStart(Text.From([Zip], "en-US"), 5, "0")

OK Cancel

123	PositionID	State	Zip	DOB	MaritalDesc	Custom.1
1	19	MA	1960	7/10/1983	Single	01960
2	27	MA	2148	5/5/1975	Married	02148
3	20	MA	1810	9/19/1988	Married	01810
4	19	MA	1886	9/27/1988	Married	01886
5	19	MA	2169	9/8/1989	Divorced	02169
6	19	MA	1844	5/22/1977	Single	01844
7	24	MA	2110	5/24/1979	Single	02110
8	19	MA	2199	2/18/1983	Widowed	02199
9	19	MA	1902	2/11/1970	Single	01902
10	14	MA	1886	1/7/1988	Divorced	01886
11	19	MA	1902	1/12/1974	Married	01902
12	19	MA	2062	2/21/1974	Married	02062
13	9	TX	78230	7/4/1988	Divorced	78230
14	19	MA	1810	7/20/1983	Single	01810
15	19	MA	2747	7/15/1977	Divorced	02747
16	19	MA	2050	10/18/1981	Single	02050

Now we are done in Power Query and move to the model. The first step is to update Salary and fully loaded cost - change type to currency.



File Home Help Table tools Column tools

Name: Salary

Data type: Whole number

Format: Whole number

Summarization: Sum

Data category: Uncategorized

Sort by column: Sort

Employee_Name	EmpID	MarriedID	Whole number	mpStatusID	DeptID	PerfScoreID	Salary	PositionID
Alagbe,Trina	10088	1	Percentage	1	5	3	64,991	15
Anderson, Carol	10069	0	Scientific	5	5	3	50,825	15
Athwal, Sam	10062	0		1	5	3	59,365	15
Bachiochi, Linda	10114	0	0	3	5	3	47,837	15
Baczinski, Rachael	10252	1	1	5	5	3	54,670	15
Barbara, Thomas	10242	1	1	5	5	3	47,211	15
Barone, Francesco A	10265	0	0	1	5	3	58,709	15
Barton, Nader	10066	0	2	5	5	3	52,505	15
Bates, Norman	10061	0	0	4	5	3	57,834	15
Becker, Scott	10277	0	0	3	5	3	53,250	15
Bernstein, Sean	10046	0	0	1	5	3	51,044	15
Biden, Lowan M	10226	0	2	1	5	3	64,919	15
Brill, Donna	10177	1	1	5	5	3	53,492	15
Bugali, Josephine	10203	0	3	3	5	3	64,375	15
Carey, Michael	10115	0	0	1	5	3	52,846	15

HR Dashboard Example - Power BI Desktop

File Home Help Table tools Column tools

Name Fully Loaded Costs \$% Format Text

123 Data type Text

Whole number

Decimal number

Fixed decimal number

Date/time

Date

Time

Text

True/false

Binary

RecruitmentSource PerformanceScore Engagem

Indeed Fully Meets

Google Search Fully Meets

Employee Referral Fully Meets

Diversity Job Fair Fully Meets

Diversity Job Fair Fully Meets

Diversity Job Fair Fully Meets

Google Search Fully Meets

On-line Web application Fully Meets

Google Search Fully Meets

0 Wednesday, March 4, 2015 0 6 668

7 804

4 660

1 770

1 677

4 754

3 809

3 788

5 792

5 701

1 719

7 696

1 640

0 Thursday, February 18, 2021 0 15 579

Data type change

With this data type change, your data will be stored differently. This may cause a loss of data or precision. After you make this change, you can restore the column by refreshing the table.

Do you want to continue?

OK Cancel