

# 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

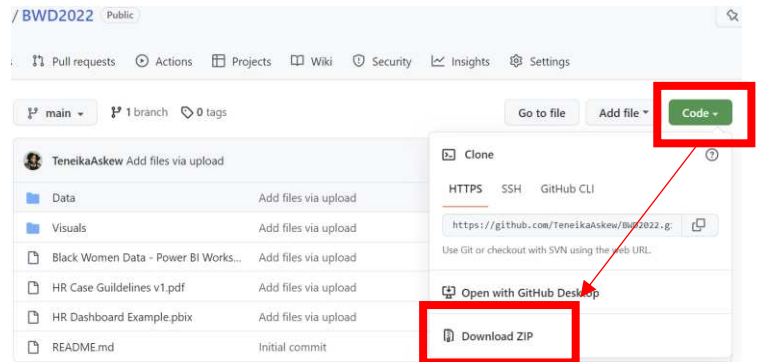
## Power BI Steps:

Outside of this workshop, you can follow the [Microsoft Power BI learning path](https://docs.microsoft.com/en-us/power-bi/learning-path/) to dive deeper into Power BI.

All files for this session are available on Github. You can download all files and relevant additional content there.

Once files are downloaded, you will need to unzip files to your Desktop for ease of use or a file path that is solely for projects so files paths are easy to configure.

<https://github.com/TeneikaAskew/BWD2022>



If you do not have a Microsoft account and want to use Power BI online or a virtual machine/lab without creating one,

here are training accounts I've created that you can access.  
<https://docs.google.com/spreadsheets/d/15ckDqvglrplvwikV-NkpA5QJ5Eb876Aa0HUK7CV-jEE/edit?usp=sharing>

### Opening Power BI Options

#### Option 1:

- Open Microsoft Virtual Machine
- Visit Become a Power BI Analyst Path
- Open up this [lab](https://learn.microsoft.com/en-us/training/modules/get-data/lab-prepare):
  - <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 to a Windows environment where Power BI is readily available

### Lab - Prepare data in Power BI Desktop

45 minutes


This unit includes a lab to complete.

Use the free resources provided in the lab to complete the exercises in this unit. You will not be charged.

Microsoft provides this lab experience and related content for educational purposes. All presented information owned by Microsoft and intended solely for learning about the covered products and services in this Microsoft module.

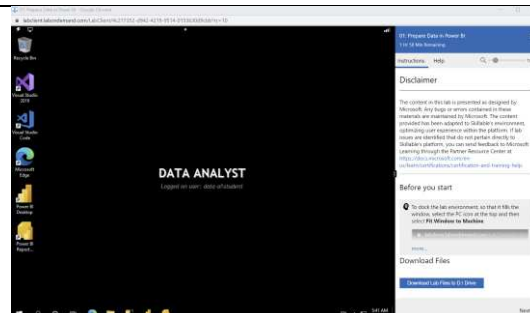
[Sign in to launch the lab](#)

### Access your environment

 Your lab environment is being built

Your lab will be ready in about 2 minutes.

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.



## Option 2:

- Download Power BI Desktop
- <https://www.microsoft.com/en-us/download/details.aspx?id=58494>
- You can download Power BI through the download option that redirects to the **Microsoft App Store** or download the executable in **Advanced download options**.



## Microsoft Power BI Desktop

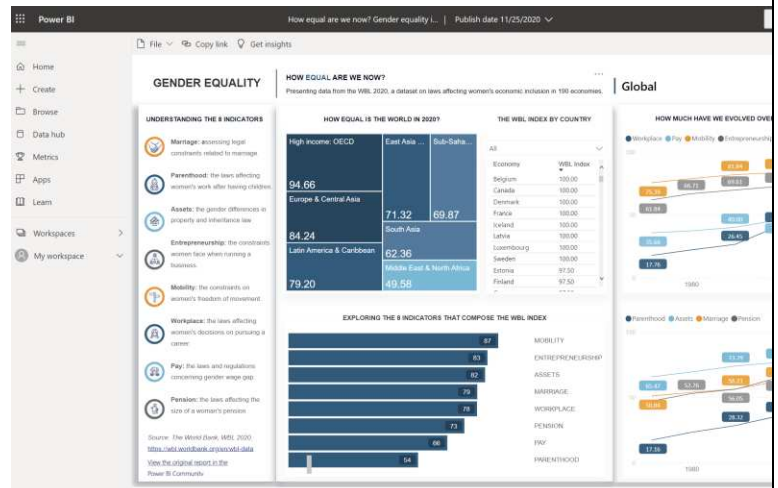
With the Power BI Desktop you can visually explore your data through a free-form drag-and-drop canvas, a broad range of modern data visualizations, and an easy-to-use report authoring experience.

**Download >**

**Advanced download options >**

## Option 3:

- Power BI online is available, but this version will look different from the Desktop version. <https://app.powerbi.com/>



## Getting Started

### 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

# Power BI Desktop





Get data



Recent sources

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

Extract Table Using Examples | Load | Transform Data

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

Queries [1]

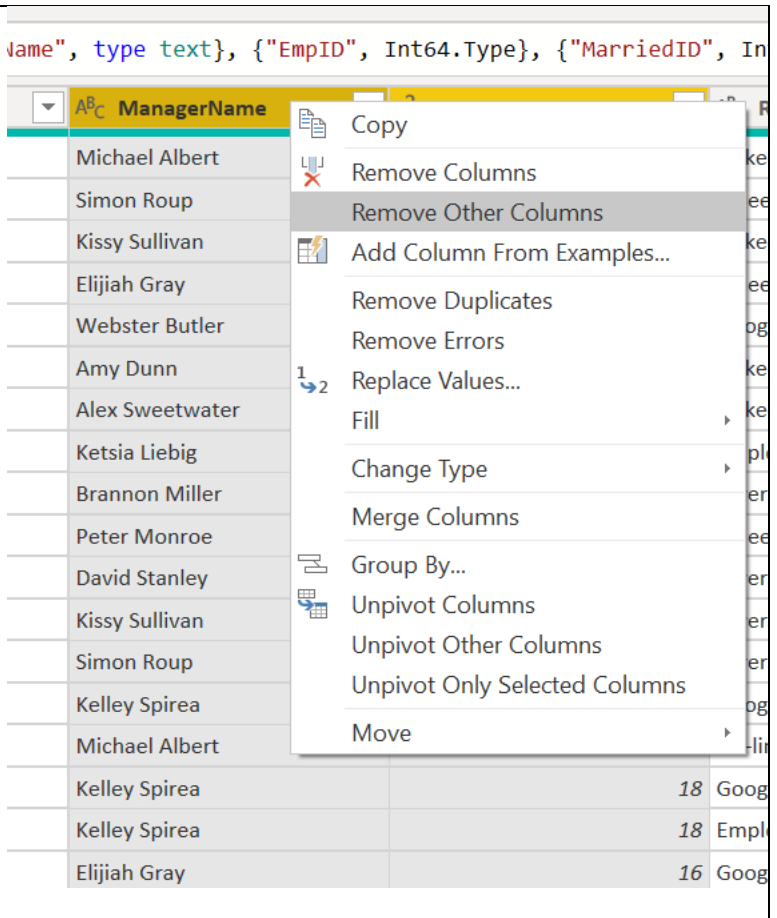

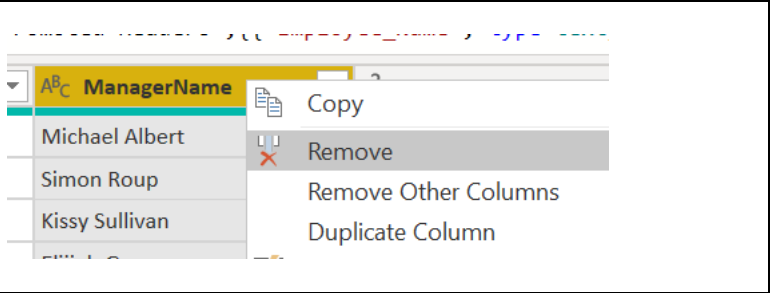
HRDataset\_v14

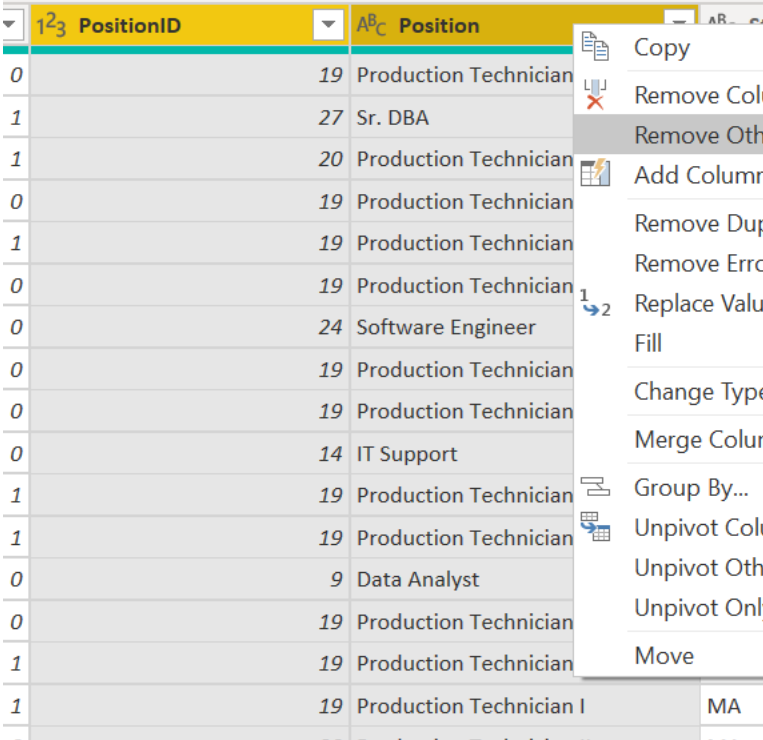
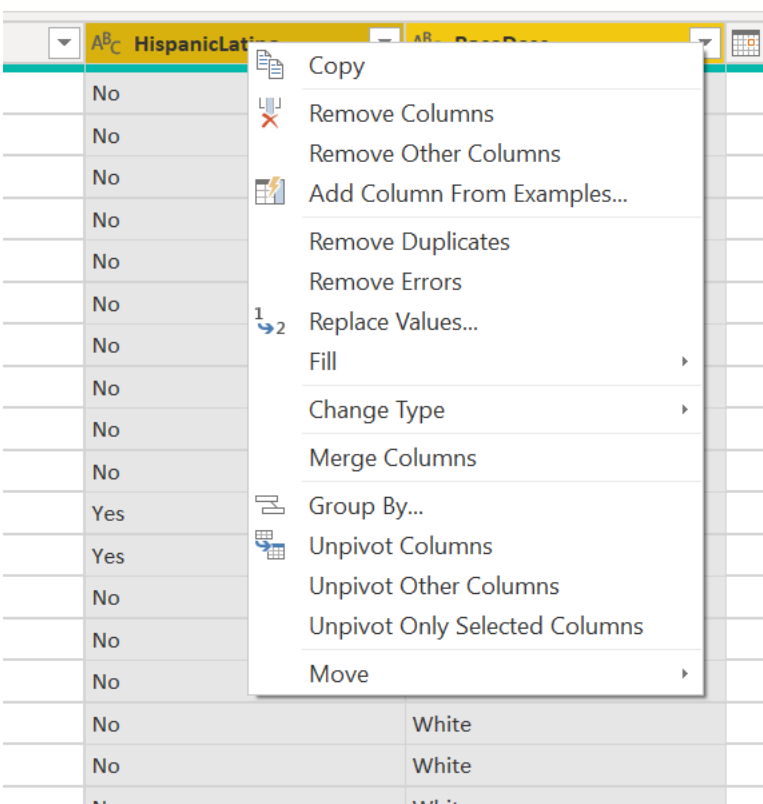
- Copy
- Paste
- Delete
- Rename
- ☒ Enable load
- ☒ Include in report refresh
- Duplicate
- Reference
- Move To Group
- Move Up
- Move Down
- Create Function...
- Convert To Parameter
- Advanced Editor
- Properties...

16 Bates, Norman

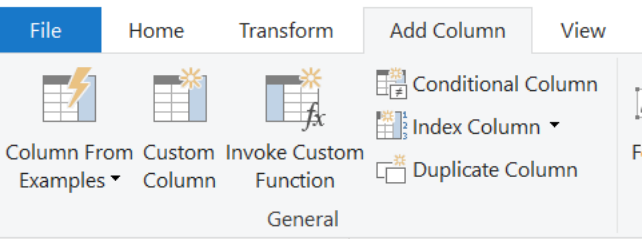
Right Click on the table and name this Manager

Remove other columns except ManagerName, Manager ID and EmpID

	
Go to the original table and rename this table from HR Dataset v14 to Employees	
Remove ManagerName from Employees	
Repeat steps to duplicate the Employees table.	
Delete Position from original data table	

	<p>headers",{{"Employee_Name", type text}, {"EmpID", Int64.Type</p> 
<p>Repeat steps to duplicate the Employees table and create Diversity table</p> <p>Delete diversity data fields from original data table</p>	<p>formColumnTypes("#Promoted Headers",{{"Employee_Name", ty</p> 
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 ⓘ  
= [Salary]\*1.25

Available columns  
Salary  
PositionID  
State  
Zip  
DOB  
MaritalDesc  
CitizenDesc

[Learn about Power Query formulas](#)

✓ No syntax errors have been detected.

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")

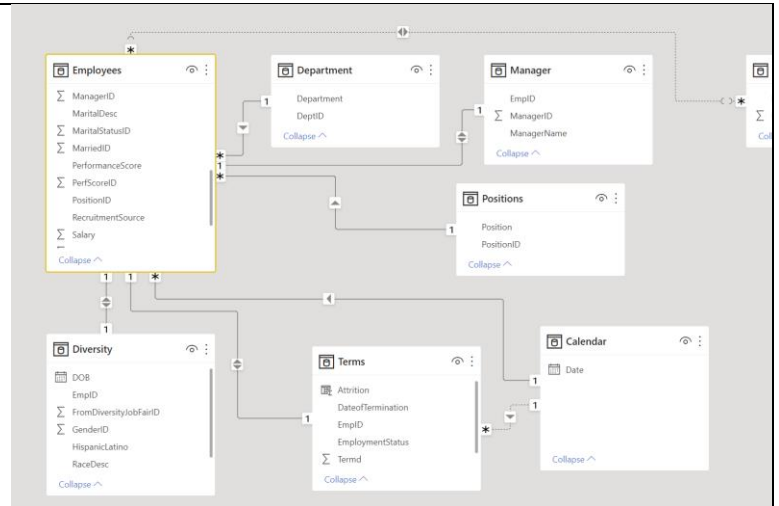
	PositionID	State	Zip	DOB	MaritalDesc	Custom.
1	19	MA	1960	7/10/1983	Single	01960
2	27	MA	2148	5/5/1975	Married	02148
3	20	MA	1810	9/19/1988	Married	01810
4	19	MA	1886	9/27/1988	Married	01886
5	19	MA	2169	9/8/1989	Divorced	02169
6	19	MA	1844	5/22/1977	Single	01844
7	24	MA	2110	5/24/1979	Single	02110
8	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

## Modeling Data

In this exercise we will not need to model the data as Power BI has autodetected the relationships.

In future dashboards you may need to create the relationships between tables.





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 Format: Whole number Summarization: Sum Data type: Whole number

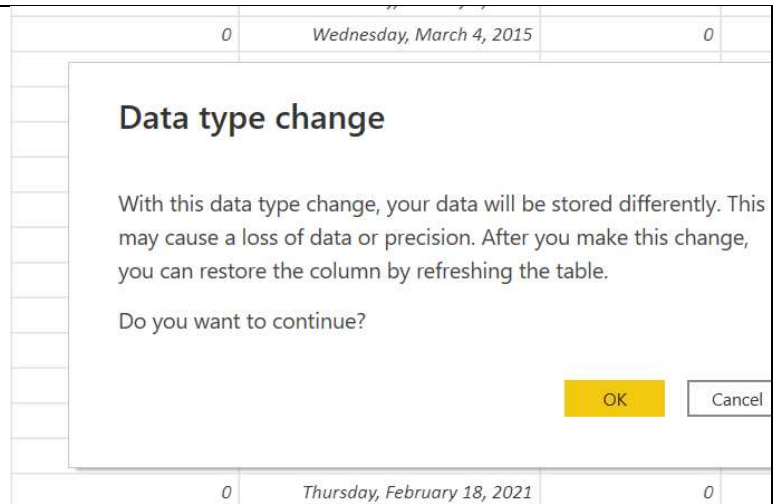
Employee_Name	EmpID	MarriedID	Whole number	mpStatusID	DeptID	PerfScoreID	Salary	Positi
Alagbe,Trina	10088	1			1	5	3	64,991
Anderson, Carol	10069	0			5	5	3	50,825
Athwal, Sam	10062	0			1	5	3	59,365
Bachiochi, Linda	10114	0		0	3	5	3	47,837
Baczinski, Rachael	10252	1	1		5	5	3	54,670
Barbara, Thomas	10242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

HR Dashboard Example - Power BI Desktop

File Home Help Table tools Column tools

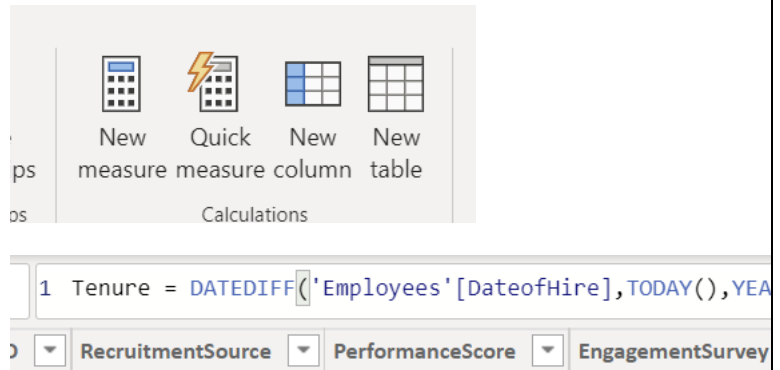
Name: Fully Loaded Costs Format: Text Data type: Text

DateofHire	RecruitmentSource	PerformanceScore	Emp
Monday, Janu	Indeed	Fully Meets	
Monday, Ju	Google Search	Fully Meets	
Monday, Septemb	Employee Referral	Fully Meets	
Monday, -	Diversity Job Fair	Fully Meets	
Monday, Janua	Diversity Job Fair	Fully Meets	
Monday, A	Diversity Job Fair	Fully Meets	
Monday, Februa	Google Search	Fully Meets	
Monday, Septemb	On-line Web application	Fully Meets	
Monday, Februa	Google Search	Fully Meets	



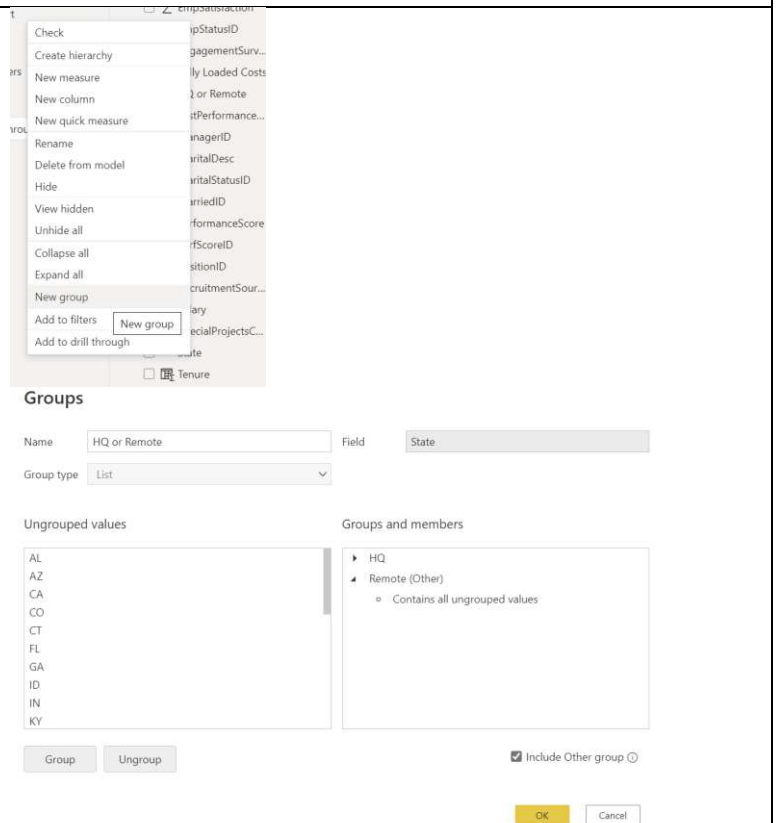
### Add a Tenure field to the model using DAX

Tenure =  
DATEDIFF('Employees'[DateofHire],TODAY(),YEAR)



### Add a HQ or Remote field to the model using Groups

- Right-click on State
- Click on **New Group**
- Click on MA then Group
- Rename the Group **HQ**
- Check **Include Other group**
- Rename the Group to **Remote**



### Add Attrition Field to **Terms** table

Attrition = **CALCULATE**(**COUNTA**(**Terms**[**Termd**]),  
**Terms**[**Termd**]=1)/**COUNTA**(**Terms**[**Termd**])

Structure

Formatting

Properties

✓

1 Attrition = CALCULATE(COUNTA(Terms[Termd]), Terms[Termd]=1)/COUNTA(Terms[Termd])

Termd	DateofTermination	TermReason	EmploymentStatus	Attrition
0172	0	N/A-StillEmployed	Active	
0127	0	N/A-StillEmployed	Active	
0174	0	N/A-StillEmployed	Active	
0135	0	N/A-StillEmployed	Active	
0010	0	N/A-StillEmployed	Active	
0043	0	N/A-StillEmployed	Active	
0271	0	N/A-StillEmployed	Active	
0084	1 Thursday, June 16, 2016	career change	Voluntarily Terminated	0.3%
0196	1 Monday, September 24,	hours	Voluntarily Terminated	0.3%
0069	1 Tuesday, September 6, 2	return to school	Voluntarily Terminated	0.3%
0252	1 Thursday, January 12, 20	Another position	Voluntarily Terminated	0.3%
0242	1 Monday, September 19,	unhappy	Voluntarily Terminated	0.3%
0066	1 Thursday, April 6, 2017	Another position	Voluntarily Terminated	0.3%
0267	1 Friday, April 4, 2014	career change	Voluntarily Terminated	0.3%

### Add a calendar table in the model using **DAX**

Calendar =  
**CALENDAR**(**FIRSTDATE**('Employees'[DateofHire]),**LASTDATE**('Employees'[DateofHire]))

Structure		Calendars	Relationships	Calculations
1 Calendar = CALENDAR(FIRSTDATE('Employees'[DateofHire]),LASTDATE('Employees'[DateofHire]))				
Date				
1/9/2006 12:00:00 AM				
1/10/2006 12:00:00 AM				
1/11/2006 12:00:00 AM				
1/12/2006 12:00:00 AM				
1/13/2006 12:00:00 AM				
1/14/2006 12:00:00 AM				

## Visualizing Data


### Select your color palette for the report

- In this example we use Classic
- You can also upload custom themes

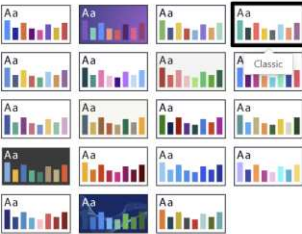
### Add a grey background

- Select the page
- Click the format icon
- Click **Wallpaper**
- Select Grey

This report



Power BI



Browse for themes  
 Theme gallery  
 Customize current theme  
 Save current theme

**Visualizations**

Format page

Search

Page information

Canvas settings

Canvas background

**Wallpaper**

Color

Image

Image fit

Transparency

Reset to default

**Fields**

Search

Calendar

Department

Diversity

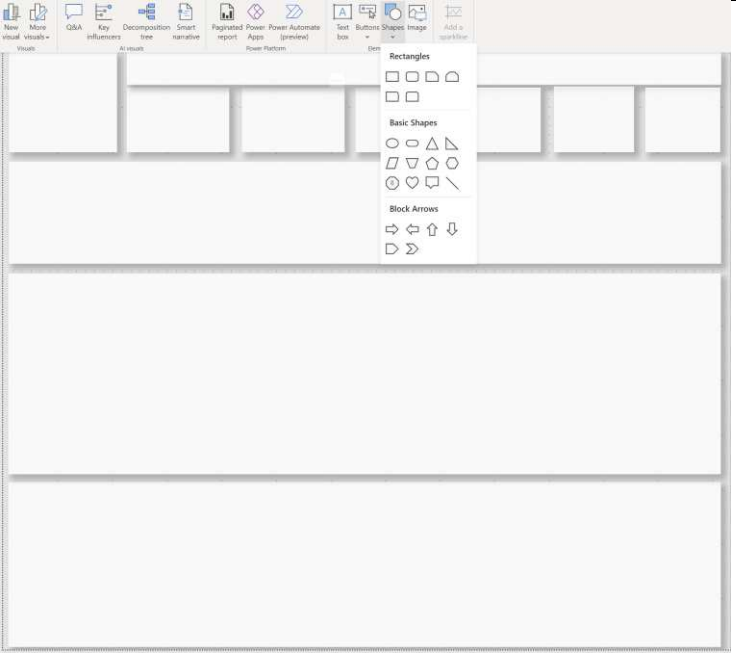
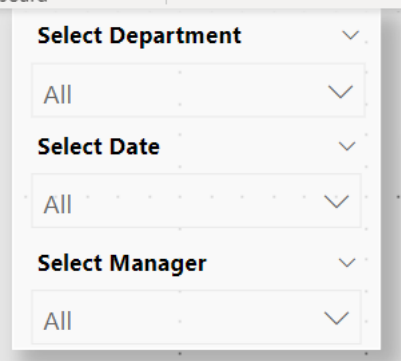

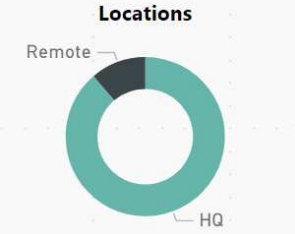
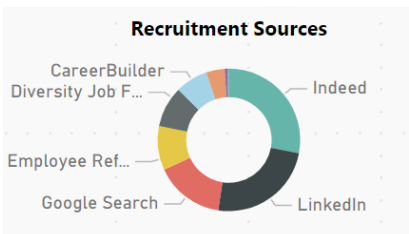
Employees

Manager

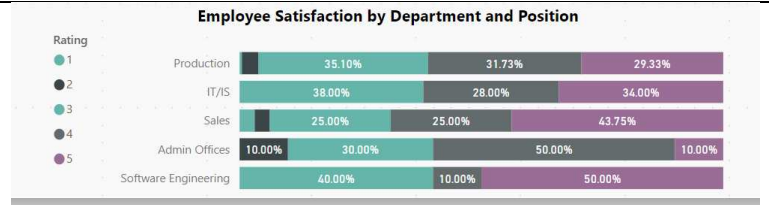
Positions

Terms

### Create your layout using the shapes option and begin to highlight your page layout of

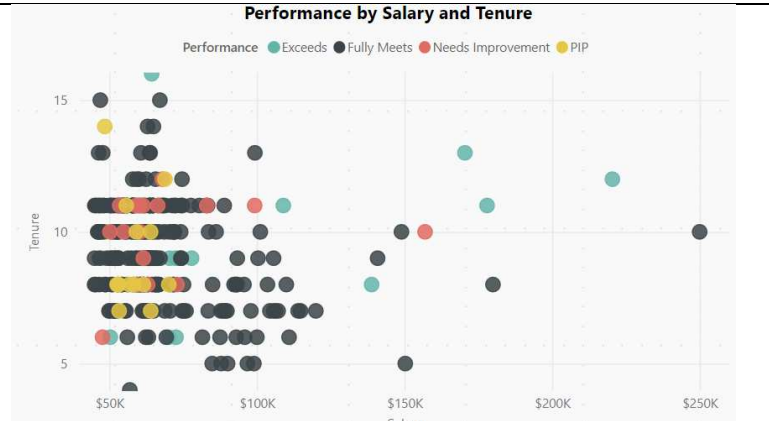
	
<p>Add a Title</p> <ul style="list-style-type: none"> <li>Select the text box and add it to the top shape</li> </ul>	<p><b>Organizational Engagement and Diversity</b></p>
<p>Slicers (Filters)</p> <ul style="list-style-type: none"> <li>Select the slicer visual</li> <li>Check Department</li> <li>Repeat for Date and Manager</li> </ul>	
<p>Hero Stats / BANs</p> <ul style="list-style-type: none"> <li>Select the Card visual</li> <li>Check Fully Loaded Costs</li> <li>Repeat for Employee ID, Absences, Tenure (Median), Avg Performance Score</li> </ul>	
<p>Location of Employees and Recruitment Sources</p> <ul style="list-style-type: none"> <li>Check Fully Loaded Costs</li> <li>Repeat and check Recruitment Source</li> </ul>	<div> <div> <p><b>Locations</b></p>  </div> <div> <p><b>Recruitment Sources</b></p>  </div> </div>
<p>Employee Satisfaction</p> <ul style="list-style-type: none"> <li>Drag Department to Y-axis</li> <li>Drag Emp ID to X-axis</li> </ul>	

- Add Performance Score to Legend
- Rename EmpSatisfaction to Rating in visual
- Exclude executive office from satisfaction. May want to exclude executive office with only 1 person in that office



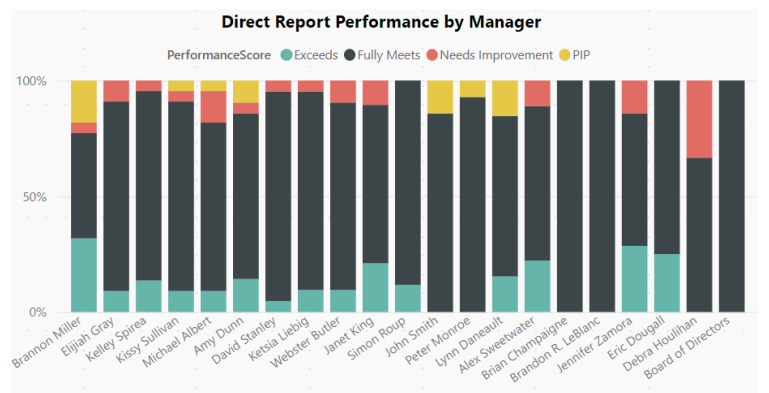
### Performance by Salary and Tenure

- Add Salary to X-axis
- Add Tenure to Y-axis
- Add Performance Score to Legend



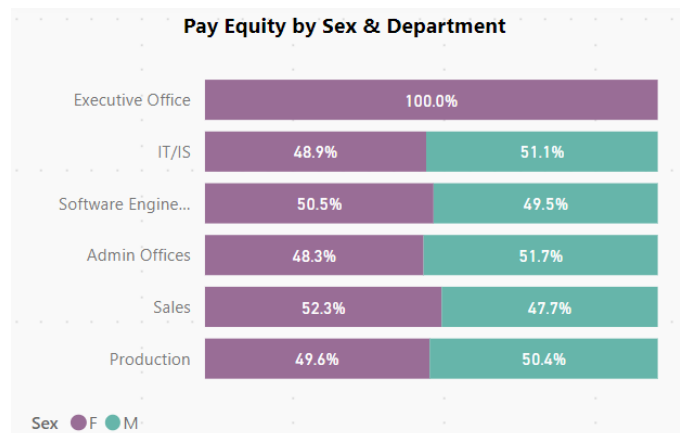
### Direct Report Performance by Manager

- Add Manager Name to X-axis
- Add Performance Score to Y-axis
- Add Performance Score to Legend



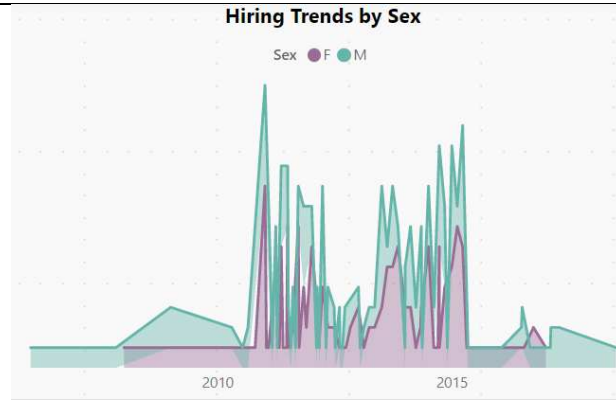
### Pay Equity by Sex and Department

- Add Department to X-axis
- Add Salary to Y-axis
- Add Sex to Legend



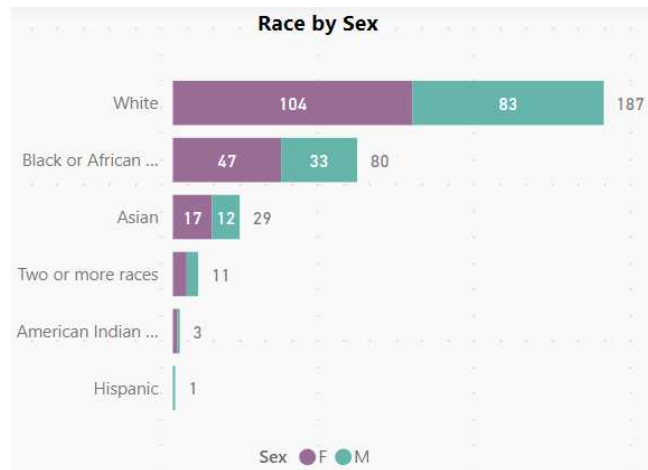
### Hiring Trends by Sex

- Add Date to Hire to X-axis
- Add Employees to Y-axis
- Add Sex to Legend



#### Diversity by Race and Sex

- Add Race Desc to X-axis
- Add Employees to Y-axis
- Add Sex to Legend



## Result



Inspiration:

<https://www.kaggle.com/code/abdelrahmanralarqan/hr-visualization/notebook>

[https://public.tableau.com/app/profile/ed.myers/viz/ExploratoryDashboard\\_16183795969740/Dashboard2](https://public.tableau.com/app/profile/ed.myers/viz/ExploratoryDashboard_16183795969740/Dashboard2)

<https://public.tableau.com/app/profile/decisive.data/viz/HumanResources-HR9Box/9BoxEmployeeRatings>

[https://public.tableau.com/app/profile/techknomatic.services.pvt.ltd./viz/AttritionDashboard\\_15576673167290/AttritionAnalysis](https://public.tableau.com/app/profile/techknomatic.services.pvt.ltd./viz/AttritionDashboard_15576673167290/AttritionAnalysis)

<https://public.tableau.com/app/profile/ajay.varghese/viz/RWFD1-HRDASHBOARD/Overview>

<https://public.tableau.com/app/profile/spencer.baucke/viz/The500HRClub/The500HRClub>