NYCT’s Utilization Analysis

**Introduction**

New York City Transit (NYCT) is a public authority, which is one of the seven agencies within the Metropolitan Transportation Authority (MTA) that operates public transportation in New York City (NYC). NYCT employs approximately 47,000 employees in various occupational groups. Under Title VII of the Civil Rights Act of 1964, the Equal Employment Opportunity Commission requires public/private employers to report the composition of its workforce by sex and race/ethnic categories. (eeoc.gov) According to the EEOC website, “Title VII of the Civil Rights Act as amended by the Equal Employment Opportunity Act of 1972 prohibits employment discrimination based on race, color, religion, sex and national origin.” In an effort to comply with EEOC guidelines private and public employers should analyze internal and external workforce data to ensure that protected employees appear in their workforce. (Note that: for the purposes of this analysis protected employees are considered Females and Minorities (i.e., Black, Hispanic, Asian, Other Pacific Islander and Hawaiian Native)). Therefore, a utilization analysis should be developed to compare the availability of females and minorities for each EEO- 4 Job Category within the five boroughs of New York City (NYC) to New York City Transit’s (NYCT) current workforce representation of females and minorities. The purpose of the utilization analysis is to determine whether underutilization of a protected class exists. Underutilization is present if NYCT’s current workforce has fewer females or minorities than would be reasonably expected by their availability in the relevant job market (NYC). When a Job Category is deemed to be underutilized, appropriate goals should be established to increase the number of females and/or minorities workers. The method used to determine underutilization is the 80% Rule.

EEO- 4 Job Categories

As mentioned above, under Title VII of the Civil Rights Act of 1964, public/private employers are required to report the composition of its workforce by sex and race/ethnic categories. The EEO-4 report is the prescribed form to be submitted to the Equal Employment Opportunity Commission (EEOC) at the end of the employer’s fiscal year. EEO- 4 Job categories are defined job functions established by the EEOC for grouping and reporting information on employees.

See below for EE0-4 Job Categories.

|  |
| --- |
| EEO-4 Job Categories |
| Officials/Administrators |
| Professionals |
| Technicians |
| Office/Clerical |
| Skilled Craft |
| Service Maintenance |
| Protective Service |
| Paraprofessional |

Figure 1: EEO-4 Job Categories

NYCT Occupational Groups

Occupational groups are the system used by NYCT to classify occupations. NYCT occupational groups are defined job functions defined by NYCT’s office of Management and Budget.

|  |  |
| --- | --- |
| **NYCT Occupational Group** | **Additional Info.** |
| Hourly Occupational | Available workforce for Operating Supervisors |
| Operating Supervisors |  |
| Other Professional & Technical |  |
| Administrative & Clerical |  |
| Managerial |  |
| Engineering |  |

80% Rule

The Office of Federal Contract Compliance Programs (OFCCP) is a part of the U.S Bureau of Labor and is tasked with ensuring employers conducting business with the Federal government are in compliance with nondiscrimination laws and regulations. The OFCCP currently recognizes four tests to determine whether underutilization exists. The four tests are: the any difference rule, the one person rule, the two standard deviation analysis and the 80% rule. NYCT’s underutilization of females and minorities will be determined by the 80% rule. Under the 80% rule, underutilization exists if the percentage of females and minorities in a particular job group within NYCT is less than 80% of the final availability percentage for NYC.

For example, Figure 3 reflects the percentage of NYCT’s Official/Clerical employees that are Hispanic. Moreover, Figure 3 also shows the percentage of Official/Clerical Hispanic workers for NYC as well as the 80% benchmark for NYC. NYCT’s workforce percentage is greater than the 4.17% benchmark for NYC, therefore; underutilization does not exist for Hispanic’s in the job category Official/Clerical.

Figure 3: Workforce Comparison

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  | NYCT Workforce | Available workforce for NYC | |
| EEO Job Group | Race | Percentage of Official/Clerical Employees | Percentage of Official/Clerical | 80% of NYC Population |
| Office/Clerical | Hispanic | 17.50% | 5.21% | 4.17% |

Figure 3: Workforce Comparison

**The Data Set**

The datasets used in this analysis is the EEO Tabulation 2006-2010 data provided by the United States Census Bureau and New York City Transit (NYCT) workforce information extracted from the PeopleSoft 9.2 database. The EEO Tabulation 2006-2010 data is derived from American Community Survey (ACS) data, which consists of estimates on the population of New York City (NYC) by Standard Occupation Classification (SOC) cross-classified by race and sex. The NYCT data consists of information on all employees working at NYCT by Sex, Ethnicity, EEO-4 Job Category and Occupational group. The EEO-4 job category column in the NYCT data set represents one or more positions having similar requirements, content, wage rates and opportunities for advancement.

Census Dataset

The EEO Tabulation data can be found in the American Fact Finder page at: <http://factfinder.census.gov/faces/nav/jsf/pages/searchresults.xhtml>. To obtain the relevant labor market statistics for NYC, first select the EEO Occupation Codes option (at the bottom of the page) and then select the SOC codes applicable to NYCT. Next, click on the “Geographies” option, select the geographic type “County - 050”, select the state NYS, choose the five counties of NYC (i.e., Kings, Queens, New York, Richmond, and Bronx). On the search result page, select the dataset named “[EEO 2r. Detailed Census Occupation by Sex, and Race/Ethnicity for Residence Geography, Total Population](javascript:;)”. The subsequent generated table includes the number of eligible individuals in the relevant labor market by SOC. Finally, download the dataset to a CSV file.

NYCT Dataset

The NYCT data can be found in the Human Capital Database. The data queried was the total number of employees in each of the eight EEO-4 Job categories by Race, Gender and internal Occupational Group classification.

Figure 2: Extract of Census Data

The Census data was bifurcated into Sex and Race groupings as minorities and females should be reflected on the Utilization Report. The percentage of workers in each EE0-4 Job category for NYCT and NYC was calculated by Race and Sex. Underutilization for Race was determined by comparing the NYCT percentages of employees in each job category, by race, to the corresponding percentages of workers by NYC’s benchmark labor market (80% of NYC labor market). Further, underutilization for females was determined by comparing the NYCT percentages of employees in each job category, by Sex, to the corresponding percentages of workers in NYC’s benchmark labor market (80% of NYC labor market).

**Officials/Administrators Census Workforce**

The available workforce for Officials/Administrators is not only the geographic location of NYC but instead a combination of NYCT data and Census data.

The EEO-4 job group Officials/Administrators used by NYCT consists of various NYCT Occupational groups.  The total NYCT employees in the EEO-4 job category Officials/Administrators are 7,802. The number of "Operating Supervisors" that are in the EEO-4 Job Group Officials/Administrators is 5,219.  As such, 67% of the Officials/Administrators population is Operating Supervisors.

As mentioned above, the available workforce for Operating Supervisors (sub-group of the Officials/Administrators population) is not the geographic location of New York City but instead the NYCT occupational group called “Hourly Occupational”. Hourly employees will take a Promotional Civil Service exam to promote to an Operating Supervisor position. Further, the number of NYCT employees in the occupational group – Hourly Occupational is 38,346. A weighted average of the two groups 67% of the Hourly Occupational group and 33% of the New York City workforce (Census data) should be used to get the total available workforce for the EEO-4 Job Group Officials/Administrators.

**Race and Sex data**

Race and Gender data was segregated into different data frames as the calculation for total NYCT race and NYC race relies on Race data (i.e., White, Hispanic, Other, Black, AI/AN, Asian, NHOPI, and Other). Moreover, the total NYCT Sex and NYC Sex relies on Sex data (i.e., Male and Female). See R Code below for steps in calculating the number and percentage of employees in NYCT and workers in NYC by Race and Sex.

**Logarithmic scales**

**According to** The logarithm log_bx for a [base](http://mathworld.wolfram.com/Base.html) b and a number x is defined to be the [inverse function](http://mathworld.wolfram.com/InverseFunction.html) of taking b to the [power](http://mathworld.wolfram.com/Power.html) x, i.e., b^x. Therefore, for any x and b,

**http://mathworld.wolfram.com/Logarithm.html**

**Analysis**

Census workforce data and NYCT workforce data were downloaded into R Studio via Read.csv and checked for missing values using na.omit.

##Check for Missing Values

#Join OCC codes to the Census data

The utilization report will compare percentages of employees in each EEO-4 Job Category, cross classified by Race and Sex. Therefore, EEO-4 Job Categories should be added to the Census data. A table with Occupational Classification Code (OCC) that is mapped to EEO-4 Job Categories was read into R Studio. Further, the Census data and the dataset that contains EEO-4 Job categories were joined by OCC. As a result of the join, a new column with EEO-4 data was appended to the Census data.

## Select geographic data on the number of residence in NYC by Sex and Race

## Rename column headers

## Sort by EEO Job Category

## Find the vertical summation for each Race and Sex category

The Census data consists of multiple statistics (e.g., totals for Race for a county within NYC for both Sexes, Margin of Error for the number of residence by Race and Sex for a particular County within NYC, etc.). The selected data is comprised of the number of residence in NYC by Race for one Sex (i.e., Male or Female). Column headers were renamed to a naming convention that would allow Race and Sex to be later separated. The following is an example of the naming convention - White.alone\_Male.

The Census data was grouped by EEO-4 job Category and then summarized by the grouping selection. The summarize function also took a vertical summation of all residence within the five boroughs by Race and Sex for each EEO-4 Job Category.

The Column headers for Race and Sex were collapsed into rows using the Gather function.

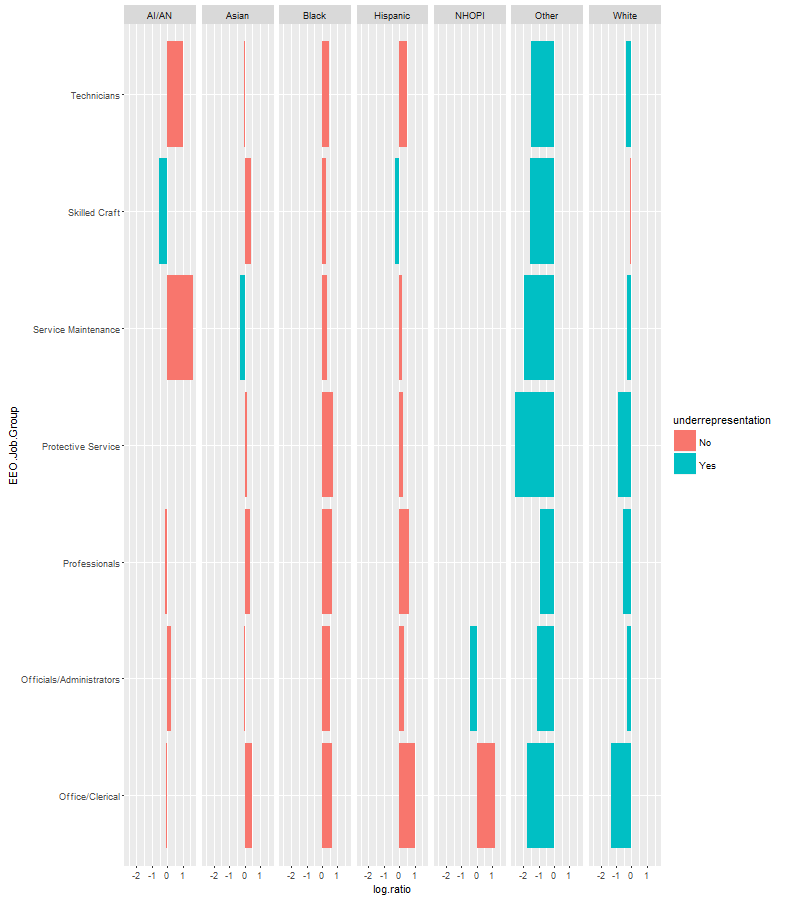
An extract of the census data before summarizing the data is provided below as Figure 2.

Figure 2

|  |  |  |
| --- | --- | --- |
| **OCC** | **occ.group** | **EEO.Job.Group** |
| 20 | General and operations managers 0020 (SOC 11-1021) | Officials/Administrators |
| 60 | Public relations and fundraising managers 0060 (SOC 11-2031) | Officials/Administrators |
| 100 | Administrative services managers 0100 (SOC 11-3011) | Officials/Administrators |
| 110 | Computer and information systems managers 0110 (SOC 11-3021) | Officials/Administrators |
| 120 | Financial managers 0120 (SOC 11-3031) | Officials/Administrators |
| 150 | Purchasing managers 0150 (SOC 11-3061) | Officials/Administrators |

**Workforce Utilization Report Results**

Under the job category Technicians, underutilization exists for American Indian/Alaskan Natives. Under the job category Skilled Craft, underutilization exists for American Indian/Alaskan Natives and Hispanics. Under the job category Service Maintenance, underutilization exists for Asians. Under the job category Professionals, underutilization exists for Females. Under the job category Officials/administrators, underutilization exists for Native Hawaiian, Other Pacific Islander and Females. Under the job category Office/Clerical, underutilization exists for Females.



sex. The NYCT data consists of information on all employees working at NYCT by Sex, Ethnicity, and Standard Occupational Classification. The job category column in the NYCT data set is one or more positions having similar requirements, content, wage rates and opportunities for advancement.

Census Dataset

The EEO Tabulation data can be found in the American Fact Finder page at: <http://factfinder.census.gov/faces/nav/jsf/pages/searchresults.xhtml>. To obtain the relevant labor market statistics for NYC, first select the EEO Occupation Codes option (at the bottom of the page) and then select the SOC codes applicable to NYCT. Next, click on the Geographies option, select the geographic type “County - 050”, select the state NYS, choose the five counties of NYC (i.e., Kings, Queens, New York, Richmond, and Bronx). On the search result page, select the dataset named “[EEO 2r. Detailed Census Occupation by Sex, and Race/Ethnicity for Residence Geography, Total Population](javascript:;)”. The subsequent generated table includes the number of eligible people in the relevant labor market in each of the eight EEO-4 Job categories. Finally, download the dataset to a CSV file.

NYCT Dataset

The NYCT data can be found in the Human Capital Database .

**Analysis**

##Check for Missing Values

Data was downloaded into R Studio via Read.csv and checked for missing values using na.omit.

#Join OCC codes to the Census data

The utilization report will compare percentages of employees in each EEO-4 Job Category, cross classified by Race and Sex. Therefore, EEO-4 Job Categories should be added to the Census data. A table with Occupational Classification Code (OCC) that is mapped to EEO-4 Job Categories was read into R Studio. Further, the Census data and the dataset that contains EEO-4 Job categories were joined by OCC. As a result of the join, a new column with EEO-4 data was appended to the Census data.

## Select geographic data on the number of residence in NYC by Sex and Race

## Rename column headers

## Sort by EEO Job Category

## Find the vertical summation for each Race and Sex category

The Census data consists of multiple statistics (e.g., totals for Race for a county within NYC for both sexes, Margin of Error for the number of residence by Race and Sex for a particular County within NYC, etc.). The selected data is comprised of the number of residence in NYC by Race and one Sex (i.e., Male or Female). Column headers were renamed to a naming convention that would allow Race and Sex to be later separated. The following is an example of the naming convention - White.alone\_Male.

The data was grouped by EEO-4 job Category and the summarized by the grouping selection. The summarization also took a vertical summation of all residence within the five boroughs by Race and Sex for each EEO-4 Job Category.

**Workforce Utilization report**

Analysis

The Census data contains two headers. The first row contains descriptive data elements and the second row has a naming convention that is a correlated with the descriptive data. Prior to loading the data into R Studio, remove the first row of the dataset so the data type of the variables will be numeric.

the American Community Survey (ACS) provided by the United States Census Bureau and NYCT Human Capital data extracted from the PeopleSoft database. The ACS data consists of estimates on the population that is 18 years or older for the five counties within NYC by Sex, Ethnicity and Education Attainment for the year of 2014. The ACS data represents the available workforce within NYC. The NYCT data consists of information on all employees working at NYCT by Sex, Ethnicity, and EEO-4 Job Category. The job category column in the NYCT data set is one or more positions having similar requirements, content, wage rates and opportunities for advancement. Each Occupational group

|  |  |  |
| --- | --- | --- |
| **OCC** | **occ.group** | **EEO.Job.Group** |
| 20 | General and operations managers 0020 (SOC 11-1021) | Officials/Administrators |
| 60 | Public relations and fundraising managers 0060 (SOC 11-2031) | Officials/Administrators |
| 100 | Administrative services managers 0100 (SOC 11-3011) | Officials/Administrators |
| 110 | Computer and information systems managers 0110 (SOC 11-3021) | Officials/Administrators |
| 120 | Financial managers 0120 (SOC 11-3031) | Officials/Administrators |
| 150 | Purchasing managers 0150 (SOC 11-3061) | Officials/Administrators |

The ACS data set was installed and downloaded into R Studio. Moreover, an API key was requested and downloaded so that the acs.fetch function can be utilized. The geo.lookup function was used to identify the five counties within New York City (Kings, Bronx, New York, Richmond, and Queens). The acs.fetch function was used to pull data on the NYC labor force data by sex, ethnicity, age, and education attainment. In addition to, the NYC labor

determine whether the percentage of their "protected" employees is equivalent to the percentages reflective of the company´s local labor market. m

The utilization analysis is a comparative analysis in which the female and minority availability for each EEO subgroup, as determined in the previous section, is compared to the current workforce representation of females and minorities. As stated earlier, the current workforce should mirror the representation of females and minorities in the job market. When subgroups have less than the percentage that is available in the job market, the group is further examined to determine whether underutilization exists. If the group is determined to be underutilized in females and/or minorities, appropriate goals are set to eliminate the underutilization

Under-utilization" is defined as "having fewer minorities or women in a particular job group than what would be reasonably expected by their availability in the relevant job market." The OFCCP and the courts currently recognize three tests to determine under-utilization:

* Any Difference Rule, in which underutilization exists if there is any difference between the availability of women or minorities compared to their percentages in the employer's actual workforce;
* Two-Standard Deviation Rule, in which the current workforce percentage of a protected class cannot vary from the availability analysis percentage by more than two standard deviations;
* 80% Rule, in which underutilization exists if the percentage of females and/or minorities in a particular job group is less than 80% of the final availability percentage.

In determining under-utilization, most employers elect to use the 80% Rule. The ratio of workforce to availability is calculated as: Ratio = (W/A) x 100.

According to the Bureau of Labor and statistics, the “Standard Occupational Classification (SOC) system is used by Federal statistical agencies to classify workers into occupational categories for the purpose of collecting, calculating, or disseminating data.”

In order to be in compliance with Equal Employment Opportunity regulations NYCT compliance A utilization analysis is intended to calculate levels of "under-utilization" of minority groups across a company´s workforce for Equal Employment Opportunity

EEO

is required to NYCT is aAs committed to Employee discrimination is

 The Utilization Analysis is a comparative analysis of the availability of females and minorities for each Equal Employment Opportunity (EEO) subgroup within the five boroughs of New York City (NYC) to New York City Transit’s (NYCT) current workforce representation of females and minorities. The purpose of the utilization analysis is to determine whether underutilization exists. Underutilization is present if NYCT’s current workforce has fewer females or minorities than would be reasonably expected by their availability in the relevant job market. When a subgroup is deemed to be underutilized, appropriate goals should be established to increase the number of females and/or minorities.

Utilization Analysis are

According to the Bureau of Labor and statistics, the “Standard Occupational Classification (SOC) system is used by Federal statistical agencies to classify workers into occupational categories for the purpose of collecting, calculating, or disseminating data.”

 Each broad occupation includes detailed occupation(s) requiring similar job duties, skills, [education](https://en.wikipedia.org/wiki/Education), or experience.

The Census and NYCT data set

When underutilization is found, placement goals should be established to increase the numbers of minorities and/or women, as needed.

Initial Data Wrangling

If subgroups have less than the percentage that is available in the job marketUnder-utilization occurs when there are fewer females or minorities in a particular job group

<https://www.eeoc.gov/laws/statutes/titlevii.cfm>

|  |
| --- |
| Error in if (length(unique(intervals)) > 1 & any(diff(scale(intervals)) < :  missing value where TRUE/FALSE needed |
|  |
| |  | | --- | | > | |

Error: ggplot2 doesn't know how to deal with data of class numeric