## Live Exercise Based on Published Story, with Downloadable Data from the Story

Summary: On March 24, 20016, CBC News posted a story based on the latest findings of the so-called sunshine list, a yearly inventory of Ontario public servants that earn more than $\$ 100,000$ a year. Other provinces like Alberta have followed Ontario's lead in not only publishing this material, but making it available in the kind of downloadable, machine-readable formats that we discussed in chapters two and three.

Traditionally, when the latest version of the Ontario list is released at the end of the fiscal year on March 31, stories tend to quickly focus on the bureaucrats who made the most. The CBC News story falls into this category, though with a wrinkle, comparing the top earners to arguably the province's most influential person, the Premier.

Deadline pressures and an incomplete knowledge of applying many of the analytical skills we've covered in this book, mean that the stories -- with a few exceptions like the Toronto Star's story about the lists gender inequity - are run-off-the-mill and predictable. However, it doesn't have to be that way. For instance, by simply downloading two years' worth of data (sadly and inexplicably, salaries posted before 2014 are unavailable in machine-readable format), you use the sorting and filtering that we learned in chapter four, to tell better stories.

## What you will learn:

1. How to download the sunshine list datasets.
2. How to create a master table.
3. How to use sorting, filtering and a bit of math, to find better stories

## Task 1: How to download the sunshine list datasets.

Go to the Ontario public sector salary disclosure page which looks like this.

| Public sector salary disclosure 2015: all sectors and seconded employees <br> Information on all public sector employees who were paid $\$ 100,000$ or more in 2015 and are subject to the Public Sector Salary Disclosure Act. |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
| This data is presented in the language in which it was provided. |  |  |  |  |  |
| Search table |  | Filter | rable by Sector |  |  |
|  | Q |  |  | - | Show/hide columns |
| Searches inciude tideten columus |  |  |  |  |  |
| Table |  |  |  |  |  |
| Last name ث | First name |  | Salary Paild | Employer | Job title |
| Weisbrodt | Jorn |  | \$207,692.37 | Toronto Festival of Arts, -- | Artistic Director |
| Parlee | Forrest |  | \$108.709.89 | Mizacs Inc. | Director, Partnerships |
| Wagner | John Clyde |  | \$168,230.79 | Toronto Festival of Arts, ... | Executive Producer |
| Yue | Jenny |  | \$114,158.53 | Colleges Ontario | Serior Manager, Finance .- |
| Macdonald | Christine |  | \$100,789.02 | Mitacs inc. | Director, Business Devel... |
| Udovic | Natasha |  | \$129,807.63 | Toronto Festival of Arts, .-. | Senior Director. Corporat... |
| Tanos | Antonia |  | \$102.222.20 | Toronto Organizing ComL.. | Senior Project Managem... |
| Webb | Paul |  | \$105,001.75 | King's University College | Professor |
| Abbas | Sadiq |  | \$107,970.20 | Algonquin College | Professor |
| Abboud | Nadim |  | S108.974.48 | Algonquin College | Professor |
| Adamyk | Mike |  | 5107,920.28 | Algonquin College | Professor |
| Agate | Jeffery M. |  | \$121,234,72 | Algonquin College | Associate Director, Stude.. |
| Al-Azzaw | Abdul |  | \$107.920.20 | Algonquin College | Professor |
| Alexander | Doug |  | \$106,084.10 | Algonquin College | Professor |
| Ali | Sabah |  | \$100.523.70 | Algonquin College | Professor |
| Allaire | Dan |  | \$114,817.15 | Algonquin College | Manager, Client Support... |
| Allan | Wilma |  | \$103.883.00 | Algonquin College | Program Quality Assuran... |
| Allen | $\operatorname{lan~D.~}$ |  | \$108,489.36 | Algonquin College | Professor |
| Almuhtadi | Wahab |  | \$110,779.56 | Algonquin College | Professor |
| 1 |  |  |  |  | - |
| 123 - Next> | Last. |  |  |  | $1 \cdot 20$ of 115533 items |
|  |  |  | Print selection | Download JSON | Download CSV |

## Select the "Download CSV" tab and save the file on your hard-drive.

Because we'll want to analyse two years' worth of data, let's do the same for the 2014 numbers, which you'll find at: https://www.ontario.ca/page/public-sector-salary-disclosure-act-disclosures-2014

## Public Sector Salary Disciosure Act: Disclosures for 2014

This page is an archive of the Public Sector Salary Disclosure for 2014. The most current Public Sector Salary Disclosure information is available at ontario.ca/salarydisclosure.

Employers subject to the Public Sector Salary Disclosure Act are required to annually disclose the names, position titles, salaries and taxable benefits of their employees who are paid $\$ 100,000$ or more in a year.

This data is presented in the language in which it was provided.


## Download the CSV file.

Opening both files reveals that they are identical with the same number of columns, headings and datatypes.


Task 2: How to create a master table.
This means that we can combine the files into a master table, allowing us to analyze for salary increases, for instance.

Open a new worksheet, copy and paste the 2015 table. Scroll the first available cell in the first column.

| 115530 | Universities | LIKıC | Jelena |
| :--- | :--- | :--- | :--- |
| 115531 | Universities | Zimmerman | Brett |
| 115532 | Universities | Zoidl | Georg |
| 115533 | Universities | Zryd | Michael |
| 115534 | Universities | Zwick | Detlev |
| 115535 |  |  |  |
| 115536 |  |  |  |
| 115537 |  |  |  |

Copy the 2014 table -- minus the headers because we already have them - place your cursor on A115535, and paste.

| 12063 | unvervues | veswe | unavar | seceream | sp,wain rux unmerxy | maxcrase rivessor |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 15530 | Universities | Zricic | Jelena | \$186,254.12 | S854.40 York University | Associate Professor |
| 15531 | Universities | Simmerman | Brett | \$117,833:28 | 570232 York University | Associate Professor |
| 15532 | Universties | Zoid | Georg | \$193,974.48 | \$1,053.44 York University | Professor |
| 15533 | Universities | 2nd | Micheel | 5133,374.84 | \$793.62 York University | Associate Dean / Associate Professor |
| 15534 | Universties | Zwick | Detiev | \$213,986.04 | \$1,086,04 York University | Associate Professor |
| 15535 | Government of Ontario- Ministries | Atmond | Margot | \$123,153.63 | \$204.96 Aboriginal Affairs | Director, Corporate Management / Directrice, gestion minis |
| 15536 | Govermment of Ontario-Ministries | Aniol | Richard | \$102,85059 | \$173.98 Aborignol Affairs | Senior Negotiator / Nä̀ gocisteur principal |
| 15537 | Government of Ontario - Ministries | Bennett | Phylus | $5114,399.48$ | \$189.54 Aboriginal Affairs | Manager, lssues Management and Media Relations / Chef, 8 |
| 155381 | finumment of Ontarin-Minitries | Piend | Anna | <10amax 91 | \$176 33 Ahncioinal Affair |  |

Your master table should have 227, 190 records, which includes the column headers.

| 227186 | Other Public Sector Employers | Landau | Stacy | $\$ 138,167.57$ |
| :--- | :--- | :--- | :--- | :--- |
| 227187 | Other Public Sector Employers | Scott | Paula | $\$ 103,944.61$ |
| 227188 | Other Public Sector Employers | Parr | Andrew | $\$ 115,000.64$ |
| 227189 | Other Public Sector Employers | Mann | Harminder Ni | $\$ 114,753.95$ |
| 227190 |  |  |  |  |
| 227191 |  |  |  |  |
| 227192 |  |  |  |  |
| 227193 |  |  |  |  |

Rename the worksheet to something like "Salaries_2014-15", and save the table as an Excel workbook, using the same label. (NOTE: When working with csv files, it is important to save them as Excel workbooks, which allows for adding multiple worksheets)

Now before we begin searching for better stories, let's add a few more columns.

Insert two columns to the right of column E, taxable benefits.


Label column F, "Total Salary", which will combine the values in the Salary Paid and Taxable Benefits columns. Label column G "Taxable Benefits as \% of Total income", and use the " + " operator to add the values in the two salaries columns.


| Review | v View | Developer Table | Salaries_2014-15.xlsx - <br> Tell me what you want to do |  |
| :---: | :---: | :---: | :---: | :---: |
|  |  |  |  | Conditional Form Formatting - Tak |
|  | D | E | F | G |
| : | Salary Paid | Taxable Benefits | Total Salary | Taxable Benefits : |
|  | \$207,692.37 | \$1,272.92 | \$208,965.29 |  |
|  | \$108,709.89 | \$0.00 | \$108,709.89 |  |
| 3 | \$168,230.79 | \$867.92 | \$169,098.71 |  |
|  | \$114,158.53 | \$131.94 | \$114,290.47 |  |
|  | \$100,789.02 | \$0.00 | \$100,789.02 |  |
|  | \$129,807.63 | \$867.92 | \$130,675.55 |  |
|  | \$102,222.20 | \$3,664.48 | \$105,886.68 |  |
|  | \$105,001.75 | \$622.01 | \$105,623.76 |  |
|  | \$107,970.20 | \$85.86 | \$108,056.06 |  |
|  | \$108,974.48 | \$129.56 | \$109,104.04 |  |
|  | \$107,920.28 | \$122.31 | \$108,042.59 |  |
|  | \$121,234.72 | \$304.47 | \$121,539.19 |  |
|  | \$107,920.20 | \$115.32 | \$108,035.52 |  |
|  | \$106,084.10 | \$129.56 | \$106,213.66 |  |
|  | \$100,523.70 | \$129.56 | \$100,653.26 |  |
|  | \$114,817.15 | \$292.97 | \$115,110.12 |  |
|  | \$103,883.00 | \$261.54 | \$104,144.54 |  |
|  | \$108,489.36 | \$85.86 | \$108,575.22 |  |

In column G, we'll calculate the taxable benefits as a percent of the total salary.


To get the percentage, we divide the taxable benefit by the total salary, and reformat the number as a percent with one decimal point ( a skill we learned in the tutorial on doing basic math like calculate percents ).


These new columns will come in especially handy for the suggested exercices a the end of this tutorial.

First order of business is filtering the table to find the Premier's salary. Activate the filter, and select the last name Wynne from the column B's drop-down menu.


You can see in the total salary column, F, that she earned $\$ 209,385.30$ and $\$ 209,374.80$ in 2014 and 2015, respectively. In essence, the same, with a slight adjustment to her taxable benefits. According to the CBC story, Kathleen Wynne salary (not including taxable benefits) was " 3,682 nd spot on the list. That means 3,681 people earned more than the boss. You can test that claim by returning to the 2015 list you downloaded, sorting the salary column in descending order, and searching for the Premier's name.

It almost checks out. She occupies the $3689^{\text {th }}$ spot, meaning there were 3,688 people ahead of her.

## Task 3: How to use sorting, filtering and a bit of math, to find better stories

Slight imprecisions notwithstanding, there is so much more than can be done with this data. For instance, because we have the salaries for two years, we can plug the table into a pivot table, select employees from a certain sector, filter for the ones who were on the payroll for both years and determine, copy the filtered pivot table, and use the paste-special tool to plug the new table into a separate worksheet.

From there, we could create a new column to calculate the percent difference, sort the result in descending order to see who got the largest increase.

Or f you wanted to so something much easier, sort column G, "Taxable Benefits as \% of Total income", in order to see who got the highest amount. In both instances, your findings could lead to interesting stories.

As well, you could filter for hospitals to see how administrators are being paid and compare their salary increases with the rank and file, who may have endured years of wage freezes. The same can be done with municipal employees, which could lead to stories about well-paid employees like paramedics, fire fighters and police officers whose salary range is below $\$ 100,000$, but find themselves on the list. This could mean that the city is footing massive overtime bills, perhaps supporting worker claims that they're short-staffed and overworked.

The possibilities are endless. And it begins with downloading two files, combining them, and creating two new columns.

