UK Gender Pay Gap Reporting 2017/8

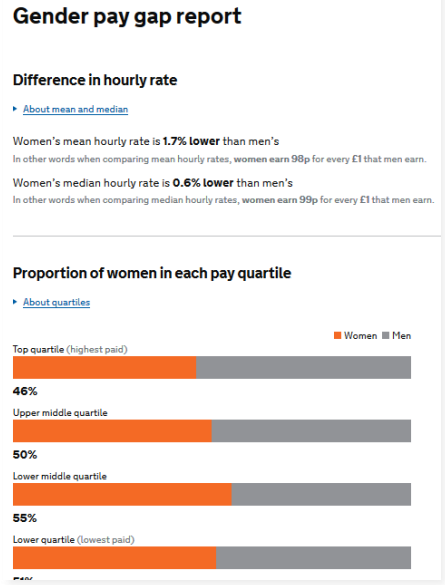
Background

Starting this year, organisations with more than 250 employees are required to report some details about the differences in pay between men and women they employ.

The government makes this data publicly available. They allow users to look up a company and see the values they reported, but provide little context for this information, or ways to see trends in aggregate data.

The visualisations constructed are aimed at two audiences: an employee wanting to see where their company stands, and journalists looking for stories either in trends or unusual cases.

gov.uk page



For Employees

The tasks likely to be most important to current employees are

- *Look up the organisation they work for*
- *Compare it to other similar organisations*

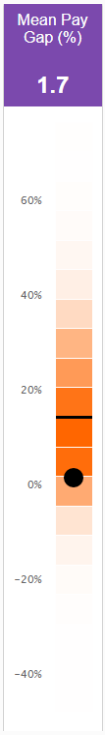
Search

A simple search box is the first thing a visitor sees, as finding their organisation is likely to be their main goal. This shows live results as they type, retrieving organisations matching their search from the server.

Profile

Clicking a company name takes a user to the profile page for that company, which contains two main elements.

Measure Bars



Each measurement we have values for is displayed in vertical bar. Most displays are wider than they are tall, so vertical bars were chosen to effectively use space.

The most important item for a first glance are the values of each measure for the company, so these are put top left, and in large bold text.

The bar is a histogram with evenly sized bins, with the transparency of colour encoding the number of companies in each bin. This lets the bar represent the distribution of companies in a small space.

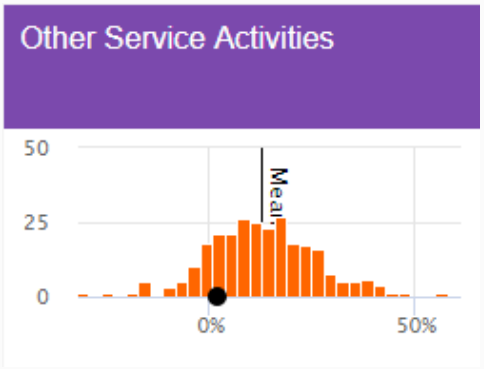
The dot is where this organisation falls, and the solid bar is the mean. The bar is not labelled due to space, but the solid bar for a mean convention is used at many other points, so is apparent from context.

Clicking on any of these bars changes the histogram display on the right side.

Histograms

The histograms on the right hand side allow comparison among areas the company works in.

These are divided up into different level of aggregation of industries defined by the Standard Industrial Classification (SIC).



A set of histograms is shown for each SIC code the organisation reported. Unfortunately, many reports did not include a SIC code, primarily from public bodies such as NHS trusts. These have all be grouped under an “Unreported” classification.

The same convention of dot for where the organisation falls and bar for mean is used for consistency

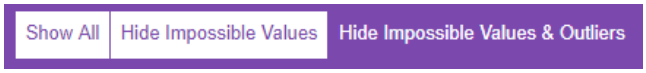
UK Gender Pay Gap Reporting 2017/8

For Journalists

A successful model for data journalism is to make data available for partners to find locally relevant stories. Examples of this model are Reveal's mortgage redlining map, or ICIJ's Panama Papers searchable network of offshore accounts.

Addresses in the data were unreliable, only reporting headquarters, so the visualisations are divided up into industrial groups. The goal is to allow journalists active in a specific field to explore and *find trends* or *find exceptions* local to that field

Outliers



An outlier could well represent an interesting story. However, outliers also obscure trends by making histograms place all value but the outliers into one bin. To satisfy both goals, all graphs are responsive to a set of always visible outlier buttons in the top right.

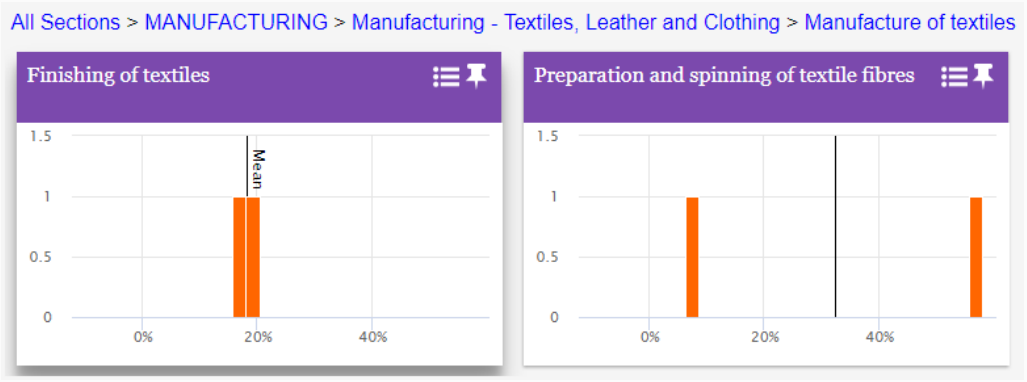
Drilldown

All the tools designed primarily for journalists allow drilling down to focus on specific industrial subgroups, to allow them to locate and visualise a specific sector. Breadcrumbs are used to maintain a sense of where in the structure is being displayed.

Explore

Each measure can be visualised in a histogram showing the distribution among the level of industry. Each level of industry can be drilled into to see it's sub groups.

The columns in these charts can be clicked to show a list of companies

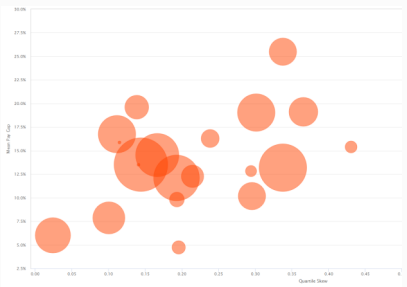


Compare

This tool allows comparison of two variables against each other, to look for relationships. Two options are provided, a scatter chart or a bubble chart.

The scatter chart plots every company in the database. As this is many thousands of points, the Highcharts boost module is used. This draws points using WebGL rather than SVG, and is considerably faster.

Clusters of points will often fall outside the general trend in the scatter graph. Identifying these groups which are performing unusually could be of interest to journalists.



The bubble chart aggregates the values for a sector and places the bubble at the mean value. Some measures, like percentage of female directors, tend to line up on similar values, making the scatter chart hard to read. The bubble chart often makes trends clearer for these measures.