Assignment 1: Descriptive and Inferential Statistics *

201374125 *University of Liverpool*

This analysis focuses on the hours worked per week and the level of education in a survey of individuals provided by the UK Data Service. From this, a conclusion is drawn regarding how overall work ethic links with the suitability of employment for graduates.

Keywords: Work Ethic; Work Suitability; Hours Worked; Graduates

---

Introduction

Selected for this analysis are the two attributes, `HoursWorked`, and `HighestQual` from the Quarterly Labour Force Survey (`qlfs`) data set provided by the UK Data Service. The aim of this analysis is to determine the relationship between graduates and non graduates, and the amount of time spent working per week. The time spent working per week is used as a method to determine overall job satisfaction and work ethic from objective quantitative data. Previous studies have found that workers satisfaction levels decline strongly with their level of education through various subjective techniques (e.g. Clark and Oswald, 1996), while others have utilised Labour Force Survey (LFS) data in other ways to determine job satisfaction and found the opposite to be true (Battu, Sloane and Belfield, 2000).

In the United Kingdom many students now see their degree as having a declining role in employment outcomes due to an overly competitive graduate market (Tomlinson, 2008), which has resulted from a rapid expansion of universities and new graduates in recent years (Chevalier and Lindley, 2006).

Univariate Description

HoursWorked:

`HoursWorked` is numeric with values from -7 to 97, representing the number of hours worked in the past week. The data is atomic, meaning some numeric values represent a character string. From `meta.data$HoursWorked`, the value -7 represents a "Non-working adult" and the value 97 represents "97 or more". This variable contains 4432 *NA* values. After removing the value -7 for non working adults, the mean value is 35.7 (data is symmetrical, skew: -0.18). The variance of Hours Worked is 207.8, with a standard deviation of 14.4. The interquartile range is 16.

HighestQual:

`meta.data$HighestQual` shows that `HighestQual` stands for Highest Qualification, categorical (factor) data with 7 levels: "Aged 70+", "Degree or equivalent", "Higher education", "GCE A Level or equiv", "GCSE grades A-C or equiv", "Other qualifications" and "No qualification". The mode average for this variable is "Degree or equivalent". The proportion misclassified for Highest

---

Qualification is 0.181, meaning that a minimum 18.1% of the survey respondents could change to result in an equal distribution of each category.

Post-Stratification

Non-response, represented by any *NA* values observed in both categories may result in a bias representation of the total population, unless correction is performed prior to analysis (Crockett and Higgins, 2011). Non-response is typically imbalanced, and will randomly vary by respondent type, based on their demographic. Trends show a declining response rate in the QLFS, which indicates a greater need for non-response consideration (Crockett and Higgins, 2011).

Correction can be obtained through weighting, known as Post-Stratification, which in this case, corrects the non-equal response of individuals based on their age group. The overall distribution of the Age Groups observed in this dataset were used to determine the likely category a non-response would belong to. Given my variables both had good response, the weightings do not affect the survey results, but are still included throughout. The worst response by age group, for example, was the 16-25 group, any response from an individual in this age group was assigned a weighting of 1.13.

Recoding

Once the weightings were obtained, the categories for `HighestQual` were recoded from 7 into 2; "Degree" and "No Degree", as my analysis concerns only the difference between graduates and non graduates. All non working adults have been excluded from the data as my analysis only concerns work ethic in relation to hours worked. `"Aged 70+"` has been excluded from either category as the data appears to exclude degree information for this age range, however most persons aged over 70 will not be working.

`HoursWorked` was first classified into levels, altering it from a numeric variable into categorical. Initially a quantile breakdown for classes was chosen, after removing the "Non-working adult" -7 value, giving levels: "0, 23, 37, 40, 45, 97" in addition to a level for "Non-working adult". These values indicated a breakdown similar to that of part-time work, full-time work and overtime work. The number of categories were reduced to reflect typical working patterns, giving: "Non working", "Part-time (0-35)", "Full-time (35-40)", "Overtime (40+)".

The UK government website gov.uk states:

> *A part-time worker is someone who works fewer hours than a full-time worker. There is no specific number of hours that makes someone full or part-time, but a full-time worker will usually work 35 hours or more a week.*

For this reason, the minimum full time hours for this analysis were selected as 35 and any persons working over 40 hours a week were considered to be working overtime.

Results

Persons who hold a degree qualification tend to work longer hours, with 52% working above 40 hours a week, compared with 42% of those who do not hold a degree (Table 1). Additionally people working part time are less likely to hold a degree qualification (25% with a degree, 35% without).

Table 1: Degree Qualification, by Hours Worked

	Degree	No Degree	All Persons
Part-time (0-35)	25	35	32
Full-time (35-40)	23	23	23
Overtime (40+)	52	42	45
Total Persons (=100%)	13945	33939	47883

Rose (2005) defines work ethic as commitment to paid work, through career outlook and non financial work commitment, both determined through qualitative interview questioning, 85% of employees with a degree agreed with the statement that their current job was part of a long term employment strategy, significantly higher than those without a degree. Battu, Sloane and Belfield (2000) used quantitative data from an older LFS dataset, using the proportion of persons in a specific job with a degree to determine how likely an individual working there is over-qualified, linking over-qualification to lower work ethic. Both these papers link job satisfaction with work ethic; similarly, my analysis determines work ethic through number of hours worked per week, with these results (Table 1; Figure 1) suggesting graduates typically have a high job satisfaction. This rejects the idea of an 'abandonment thesis' (Clark and Oswald, 1996) and implies that graduate job opportunities are not as low as found in various papers (e.g. Tomlinson, 2008). In order to confirm that many graduates are in work for which they are not over-qualified, a further variable from the QLFS dataset must be used (See below).
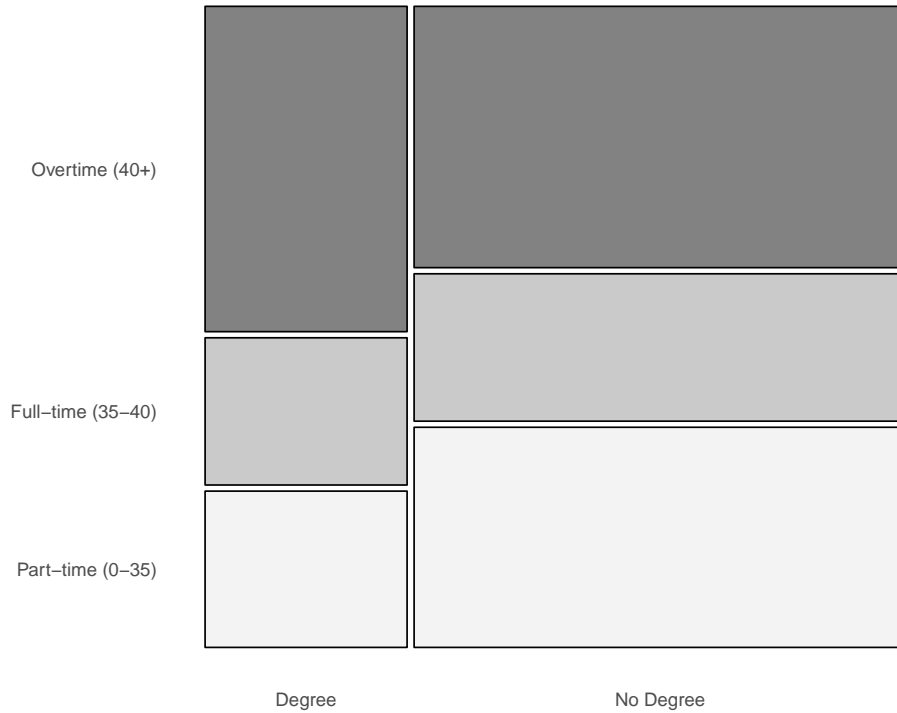

Figure 1: Degree Qualification, by Hours Worked

A significant proportion of all persons work over 40 hours a week (Figure 1), with very few working within what is generally considered to be typical hours for full time work. Despite various changes to legislation, there remains the culture that encourages long working hours across multiple disciplines throughout the UK (Barnard, Deakin and Hobbs, 2003).

Chi-squared Test

There is a statistically significant association between hours worked per week and degree qualification.

$$(X^2 = 512.05; p = 0.000)$$

This survey takes a small sample from the overall UK population, meaning this observed distribution of variables may only be due to sampling noise. The Chi Squared test is implemented in order to determine if this observed distribution is due to random sample noise, or if this sample reflects that of the whole UK population. In this case, the distribution of number of hours worked, in relation to the qualification of a respondent. The significant test result here shows that our observed distribution is not due to random noise. The test however does not indicate any relationship or strength of relationship between variables and, due to the relatively large number of samples used in this data set, the Chi Square test was very likely to give a significant statistic (Mchugh, 2013). Therefore, an additional variable is added below, in order to determine the practical significance of the relationship between these two variables.

Confidence Intervals

Confidence intervals are implemented into the visualisation of the results below to indicate the likely range of the given data value, and the percentage confidence for that range (Cumming and Finch, 2005). The Confidence Intervals used in Figure 2 below indicate a confidence of 95%. The dotted lines show that while the percentage for each category in hours worked is inversely proportional for "Degree" and "No Degree", due to these being the only assigned categories, the confidence interval varies between them. The larger confidence interval, observed in people who work full time with a degree, against those who work full time without a degree, is primarily due to the lower sample size of this variable.

The large number of respondents used to produce Figure 2 means that the confidence intervals are relatively short and that there is no overlap between the confidence intervals of any variable. We can be at least 95% confident that a significantly higher proportion of people with a degree work overtime more than they work part time or full time, with the inverse being true for those without a degree.

Adding "Type of Work"

Figure 3 adds a third dimension to the data, with the addition of the category `qlfs$NSSEC` which categorises the type of work for each respondent. Again new weightings were obtained through Post-Stratification to provide a better representation of respondents who chose not to answer their type of work. This variable has been recoded from 9 into just 2, "Professional"" and "Non-Professional" work, non working adults are again excluded. Professional work reflects a career that typically would require specific qualifications, such as a degree.

Given this data has fewer valid respondents per category, the Confidence Intervals are wider than the graph above. Where the confidence intervals overlap, the data could potentially be the
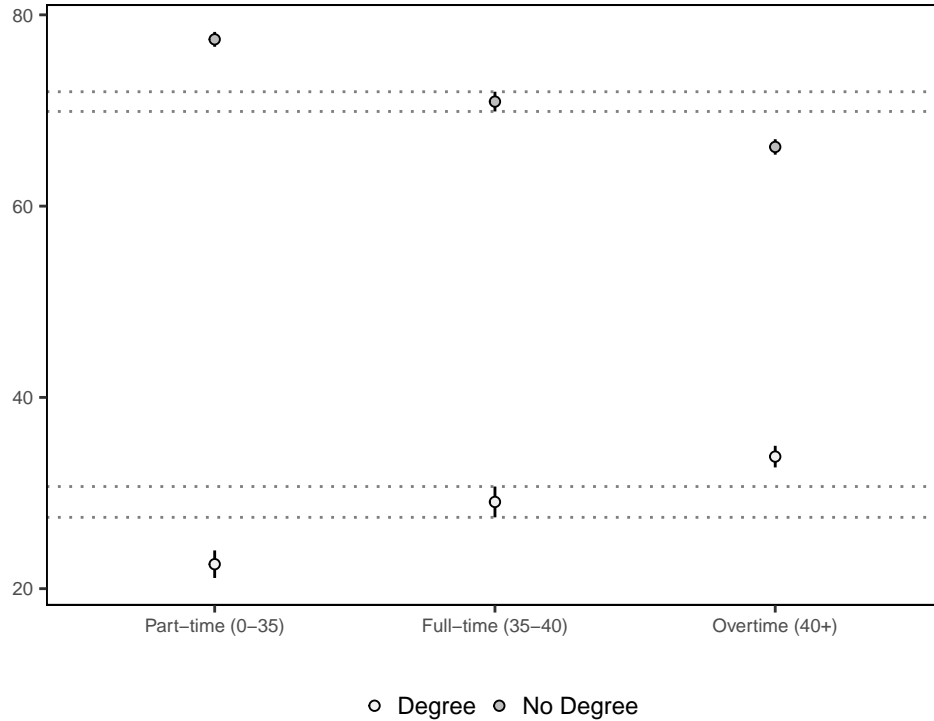
Figure 2: Confidence Intervals of Hours Worked, by Degree Qualification.

Table 2: Degree Qualification, by Job Type

	Degree	No Degree	All Persons
Professional	74	25	38
Non-Professional	26	75	62
Total Persons (=100%)	15970	44118	60088

same value. This isn't to say that there isn't a significant difference between the two values however, as for example, with Professional work and full time hours, the intervals between those with a degree and those without has a very low proportional overlap. Cumming and Finch (2005) state that for two values that have confidence intervals that just touch end to end, in a large set of data, the p value is around 0.006, far lower than the p = 0.05 that would suggest a statistically significant difference.

Figure 3 shows that a far higher proportion of people who work in non-professional jobs work part-time, and that there is no significant difference between those who hold a degree and those who do not. For full time hours there is no difference between professional and non-professional jobs. A far higher number of people who work over 40 hours a week work in a professional job, and since there is no overlap in confidence intervals between people with a degree and without a degree, a significantly higher proportion of people with a degree work overtime hours in a professional career than those who do not have a degree.

Graduates typically develop the skills that are required for many professional careers (Delom et al., 1998), as such, it is to be expected that many graduates have selected a degree in which their

5

interest lies, subsequently moving into a career that is relevant to their degree with a high work ethic. This result suggests a higher work satisfaction, and work ethic for graduates in professional work (Rose, 2005).

Lower job satisfaction in graduates may arise from over-qualification in work (Battu, Sloane and Belfield, 2000; Cabral Vieira, 2005) and studies have found this to be true, as expectation from recent graduates go unrealised in work that they are overqualified for (Tsang, Rumberger and Levin, 1991). However, Battu, Sloane and Belfield (2000) found that generally graduate work satisfaction is high, with results suggesting that many graduates avoid work for which they are over-educated. Table 2 reflects this, as 74% of respondents with a degree work in a professional career, combined with the proportionally higher number of graduates working overtime in professional work (Figure 3).
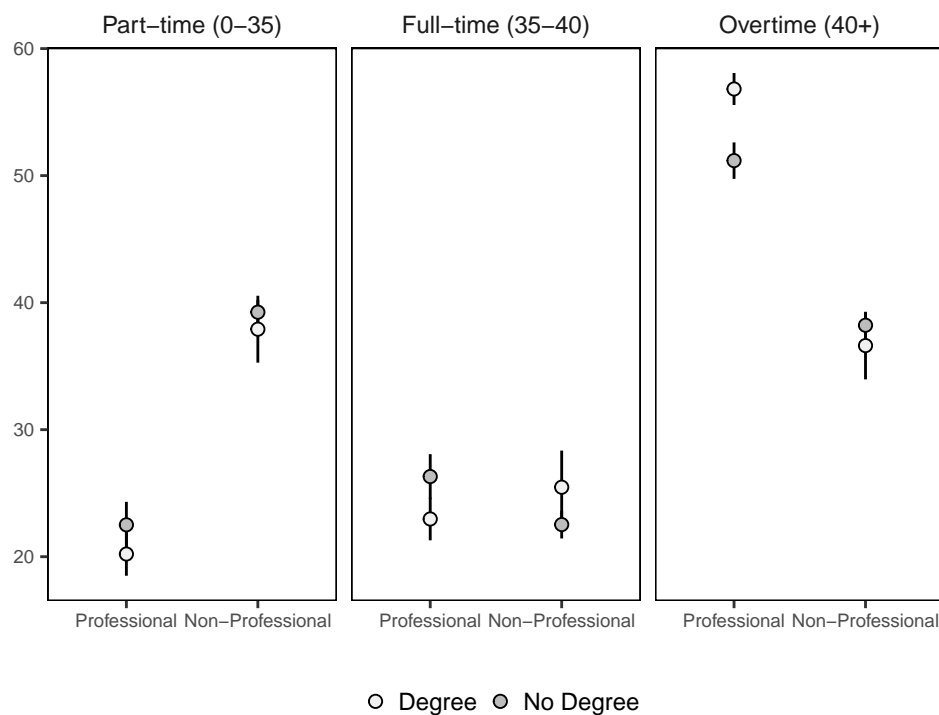

Figure 3: Type of Work, by Degree Qualification and Hours Worked

The suggestion that an increasing number of graduates has lead to an over-saturated graduate job market (e.g. Delom et al., 1998; Tomlinson, 2008), is refuted by Battu, Sloane and Belfield (2000) who find that, due to an influx of new graduates in recent years, companies have taken advantage, giving more opportunity for roles with greater responsibility and initiative, ameliorating the effects of dissatisfaction on over-qualified workers, and benefiting the company. Although there is some level of over-education in the UK, true over-education only arises from skill mismatch or lack of skills in individuals (Chevalier and Lindley, 2006).

References

Barnard, Catherine, Simon Deakin and Richard Hobbs. 2003. "Opting out of the 48-hour week: Employer necessity or individual choice? An empirical study of the operation of article 18(1)(b) of the working time directive in the UK." *Industrial Law Journal* 32(4):223–252.

Battu, H., P. J. Sloane and C. R. Belfield. 2000. "How well can we measure graduate over-education and its effects?" *National Institute Economic Review* 171(1):82–93.

Cabral Vieira, José A. 2005. "Skill mismatches and job satisfaction." *Economics Letters* 89(1):39–47.

Chevalier, Arnaud and Joanne Lindley. 2006. "Over-Education and the Skills of UK Graduates." (August).

Clark, Andrew E. and Andrew J. Oswald. 1996. "Satisfaction and comparison income." *Journal of Public Economics* 61(3):359–381.

Crockett, Alasdair and Vanessa Higgins. 2011. "Weighting the Social Surveys ESDS Government." (May):1–27.

Cumming, Geoff and Sue Finch. 2005. "Inference by eye confidence intervals and how to read pictures of data." *American Psychologist* 60(2):170–180.

Delom, Frédéric, Delphine Fessart, Marie-Elaine Caruso and Eric Chevet. 1998. *Key Qualifications in Work and Education*. Vol. 352 Dordrecht: Springer Netherlands.

Mchugh, Mary L. 2013. "The Chi-square test of independence Lessons in biostatistics." *Biochemia Medica* 23(2):143–9.

Rose, Michael. 2005. "Do rising levels of qualification alter work ethic, work orientation and organizational commitment for the worse? evidence from the UK, 1985-2001." *Journal of Education and Work* 18(2):131–164.

Tomlinson, Michael. 2008. "'The degree is not enough': Students' perceptions of the role of higher education credentials for graduate work and employability." *British Journal of Sociology of Education* 29(1):49–61.

Tsang, Mun C., Russell W. Rumberger and Henry M. Levin. 1991. "The Impact of Surplus Schooling on Worker Productivity." *Industrial Relations* 30(2):209–228.