



Baseline Data: Creating a dataset to inform our COVID-19 model

Nik Lomax, Karyn Morrissey, Jamil Nur

Overview

Our team has three objectives:

- I) Create a synthetic population at a fine geographic scale
- II) Match wider socioeconomic, health and behavioural data to create a rich baseline population individual information to our base demography
- III) Understand group and spatial patterns to inform the behavioral model



URBAN ANALYTICS / WEDNESDAY, 29 MAY, 2019



SPENSER Synthetic Population Estimation and Scenario Projection Model

Gonzalo Cruz Garcia, Nik Lomax, Andrew Smith – University of Leeds

SPENSER provides the framework for estimates of population which are dynamic and high resolution (at household level), and a comprehensive set of tools for user customisable scenario projections.

A little bit of detail

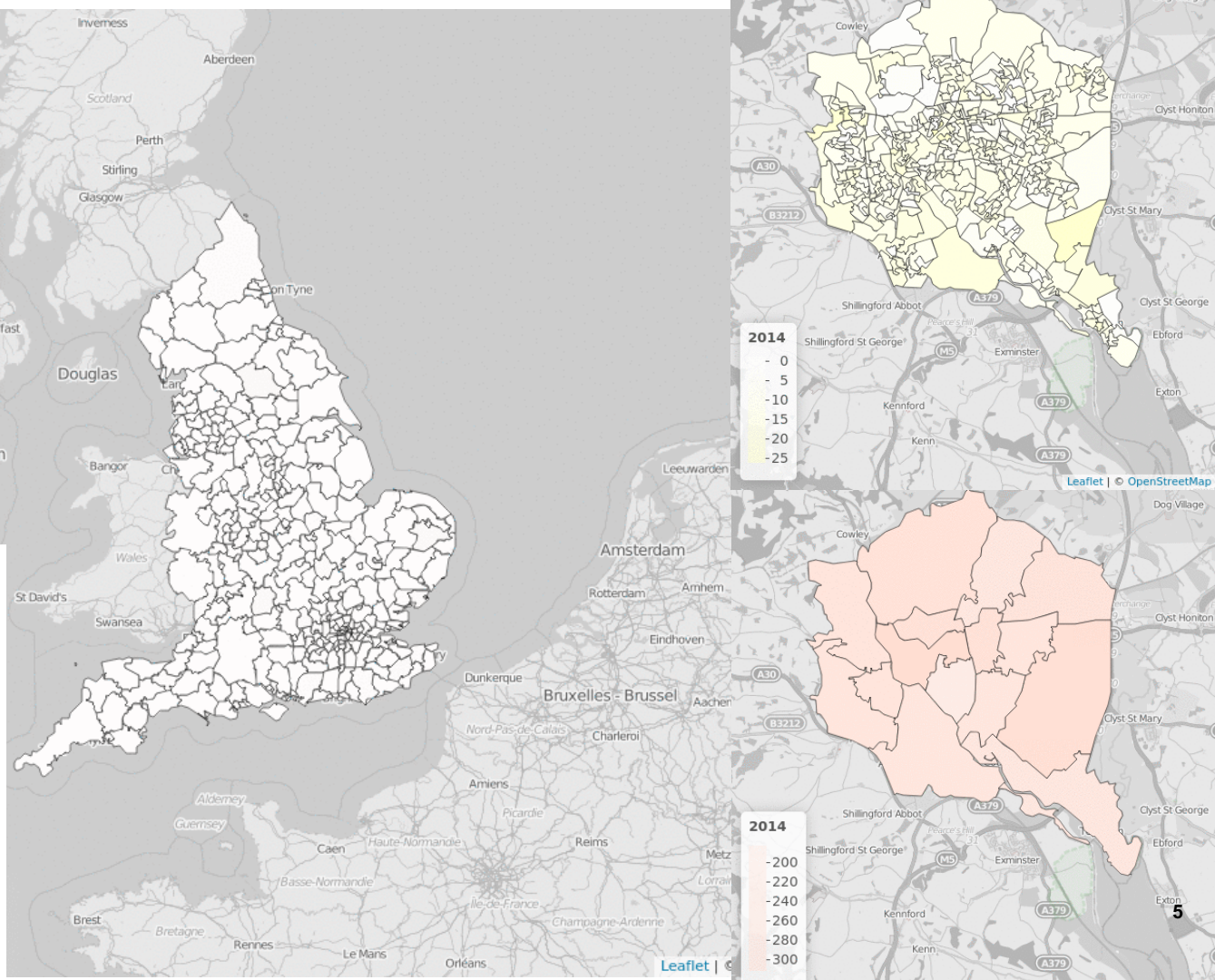
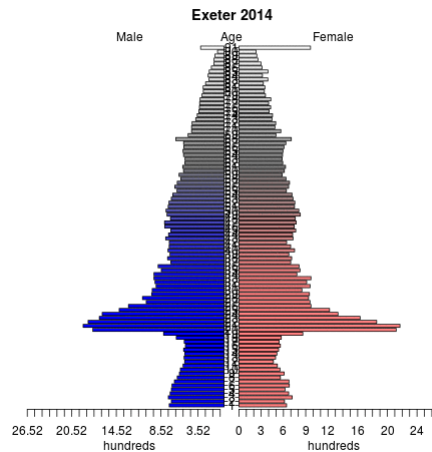
- Core implementation in C++.
- Generates synthetic populations from two or more marginal constraints
- Fast implementation of Iterative Proportional Fitting (IPF)
- Slower quasi-random sampling algorithm (QIS) which generate integer populations.
- Spatial Microsimulation package for R and python

+

-

2012-2011 household density change

-0.990%
-0.995%
-1.000%
-1.005%
-1.010%
-1.015%
-1.020%
-1.025%
-1.030%



Tracking individual status and activities

The synthetic population is our stepping-stone

Next key question: what further information do we need?

Health status: Health Survey of England (2018; 10,000 respondents)

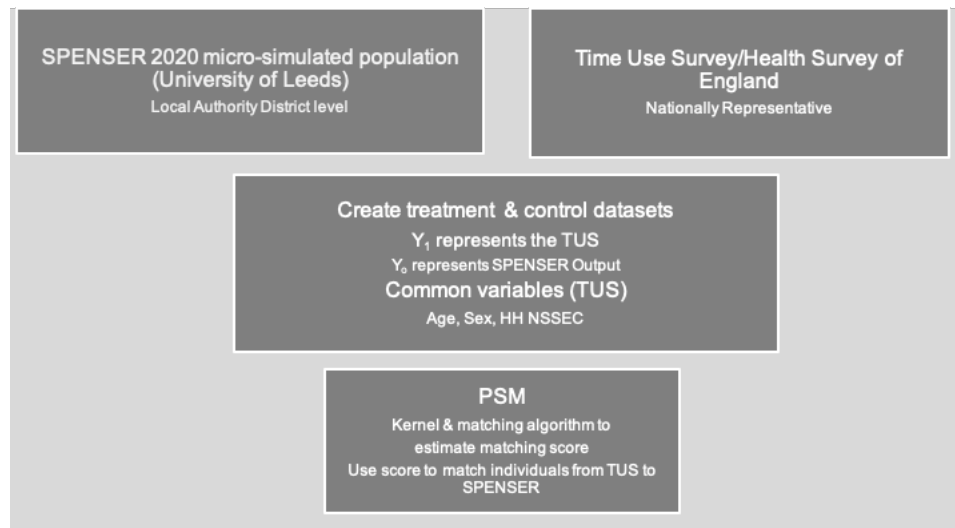
Day activities: UK Time Use Survey (2014/15; 16,000 time diaries)

Employment status: UK Time Use Survey (2014/15)

Following step: create a unique individual database at MSOA scale

Propensity Score Matching

PSM uses a calculated **‘Propensity Score’** to match individuals from a treatment dataset to a control dataset using common variables to both datasets



	area	hid	pid	Sex	age	Origin	cvd	diabetes	bloodpress-e	nssec5	phome	pwork	pschool	pshop
1	E02004446	17	2796	1	66	1	0	0	0	3. Small empl. & own account wkrs	.5555556	.	.	.
2	E02004446	17	6276	0	65	1	0	1	1	2. Intermed. occup.	.6111111	.	.	.0347222
3	E02004446	30	3211	1	75	1	0	0	0	1. Mngt, admin & prof. occup.	.8541667	.	.	.
4	E02004446	30	6385	0	67	1	0	0	0	2. Intermed. occup.	.9305556	.	.	.
5	E02004446	41	3480	1	86	1	0	0	1	5. (Semi)routine occup.	.8333333	.	.	.0625
6	E02004446	85	2566	1	62	1	1	1	1	3. Small empl. & own account wkrs	.5555556	.	.	.
7	E02004446	85	6020	0	60	1	0	0	0	2. Intermed. occup.	.6111111	.	.	.0347222
8	E02004446	109	383	1	11	1	0	0	0	0. Not applicable	.6875	.	.	.
9	E02004446	109	3337	1	79	1	0	0	0	1. Mngt, admin & prof. occup.	.8541667	.	.	.
10	E02004446	109	6482	0	69	1	0	0	0	2. Intermed. occup.	.9305556	.	.	.
11	E02004446	127	4981	0	43	1	0	0	0	1. Mngt, admin & prof. occup.	.7222222	.	.	.
12	E02004446	129	4664	0	35	1	0	0	0	2. Intermed. occup.	.5902778	.3125	.	.
13	E02004446	135	2852	1	67	1	0	1	1	5. (Semi)routine occup.	.8263889	.	.	.
14	E02004446	135	6511	0	70	1	0	0	0	5. (Semi)routine occup.	.3263889	.	.	.
15	E02004446	138	579	1	15	1	0	0	0	0. Not applicable	.6875	.	.	.
16	E02004446	138	3408	1	82	1	0	0	0	1. Mngt, admin & prof. occup.	.8541667	.	.	.
17	E02004446	138	6817	0	76	1	1	0	1	1. Mngt, admin & prof. occup.	.9930556	.	.	.
18	E02004446	179	1244	1	37	1	0	0	0	5. (Semi)routine occup.	.5138889	.375	.	.
19	E02004446	189	1805	1	50	1	1	0	1	1. Mngt, admin & prof. occup.	.6527778	.	.	.
20	E02004446	189	4151	0	16	1	0	0	0	0. Not applicable	1	.	.	.
21	E02004446	189	5085	0	45	1	0	0	0	2. Intermed. occup.	.5347222	.375	.	.
22	E02004446	223	72	1	3	1	0	0	0	0. Not applicable	.5763889	.	.2777778	.
23	E02004446	223	1869	1	51	1	0	0	0	3. Small empl. & own account wkrs	.3958333	.3611111	.	.
24	E02004446	223	5383	0	50	1	0	0	0	3. Small empl. & own account wkrs	.5763889	.	.	.0416667
25	E02004446	240	2147	1	54	3	1	0	1	5. (Semi)routine occup.	.5	.2847222	.	.
26	E02004446	240	5041	0	44	1	0	0	0	5. (Semi)routine occup.	.6111111	.	.	.
27	E02004446	244	3338	1	79	1	0	0	0	1. Mngt, admin & prof. occup.	.8541667	.	.	.
28	E02004446	244	3932	0	12	1	0	0	0	0. Not applicable	1	.	.	.
29	E02004446	244	6614	0	72	1	0	0	0	2. Intermed. occup.	.9305556	.	.	.
30	E02004446	267	1592	1	46	1	0	0	0	3. Small empl. & own account wkrs	.3958333	.3611111	.	.

Vars: 14 of 52 Order: Modified

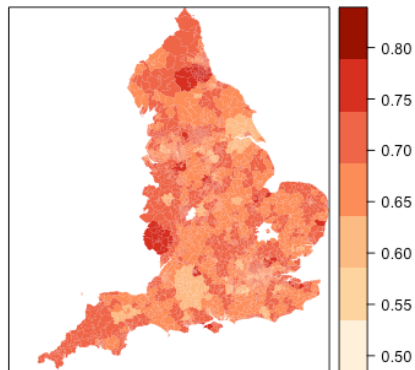
Obs: 134,562

Length: 9 Filter: Off

Detailed demographic, socioeconomic, health and time use at the individual level by household by MSOA

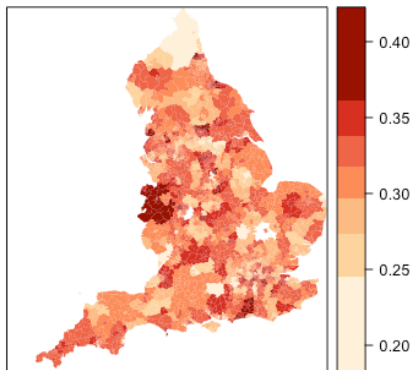
Spatial Outputs England

Average Time Spent at Home



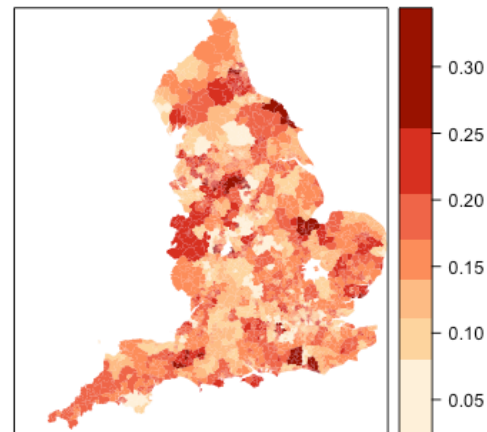
MSOA

Average Time Spent at Work



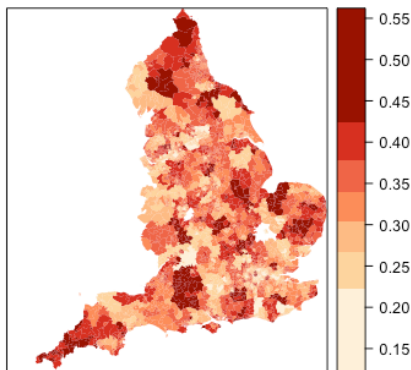
MSOA

pc_keyworker



MSOA

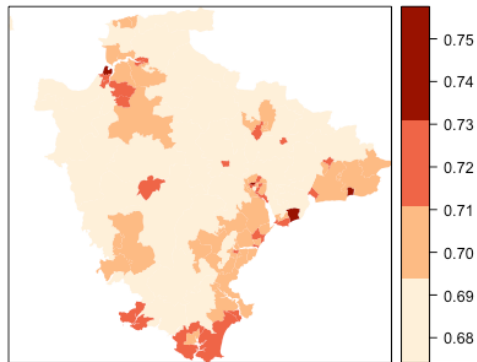
Average Underlying Health Conditions



MSOA

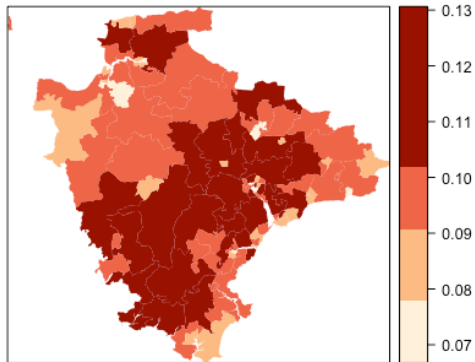
Flexible Spatial Modelling: Devon

Proportion Time Spent at Home



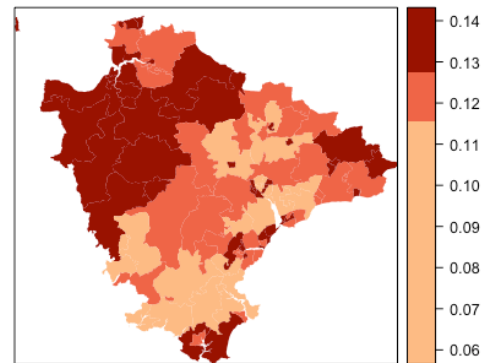
Av 70% Devon

Proportion Time Spent at Work



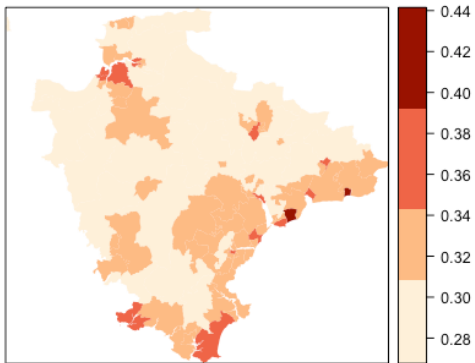
Avg 10% Devon

% keyworker



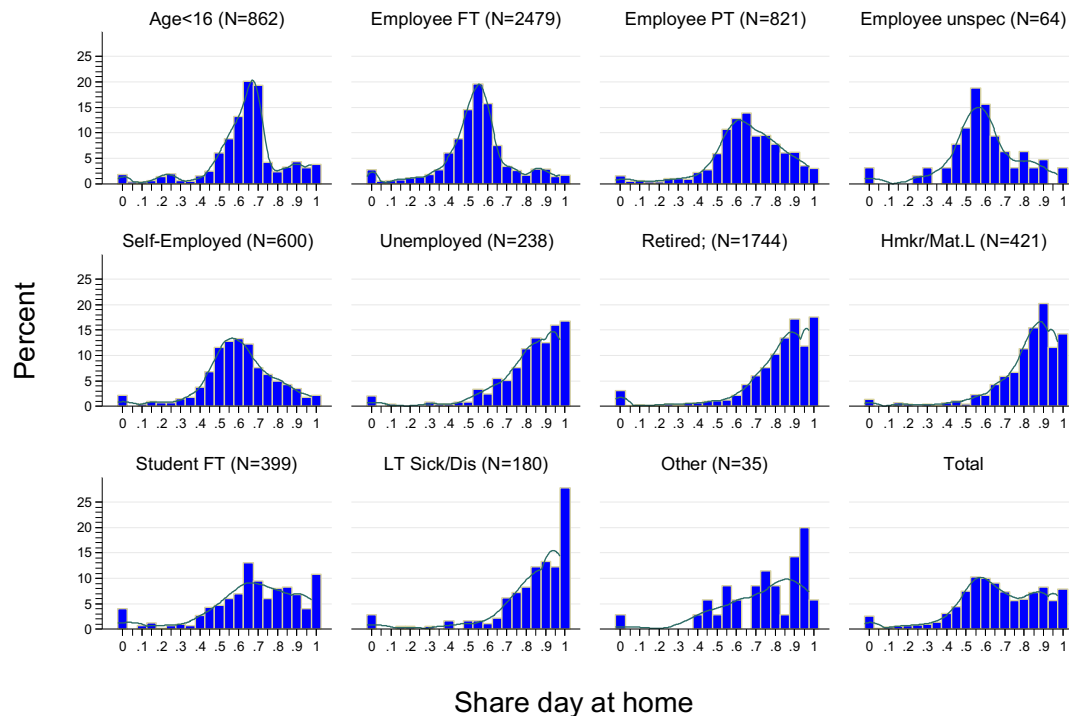
Avg 12% Devon

% Underlining Health Conditions



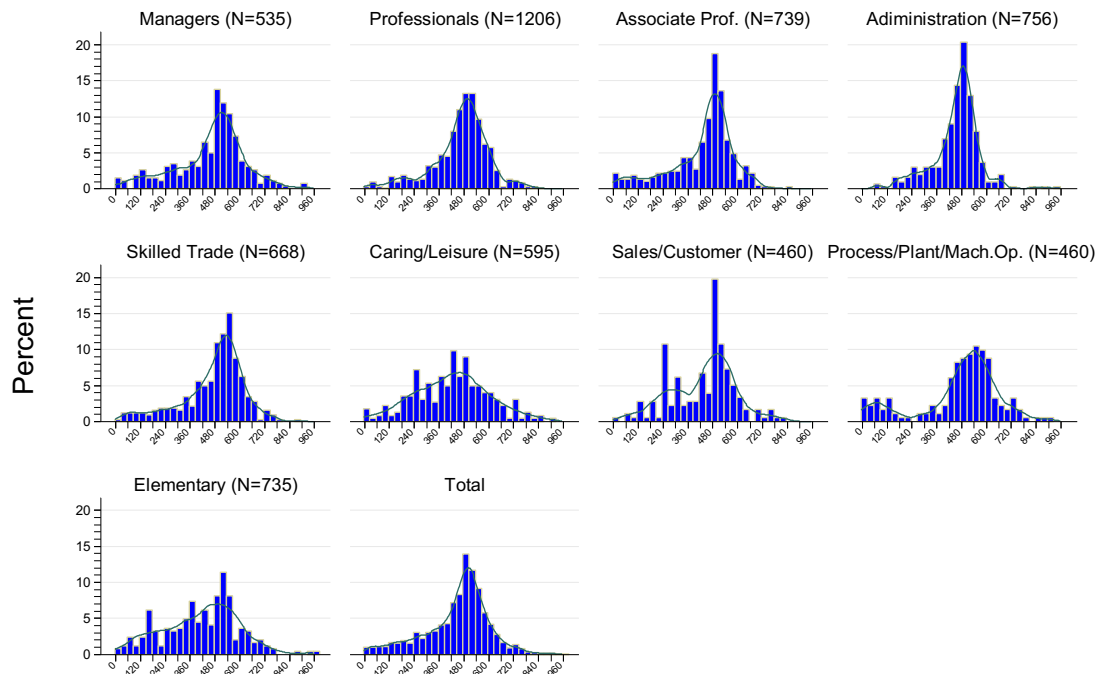
Devon Avg 32%

Time at home varies by *working status*



Total observations: 7843 person-day

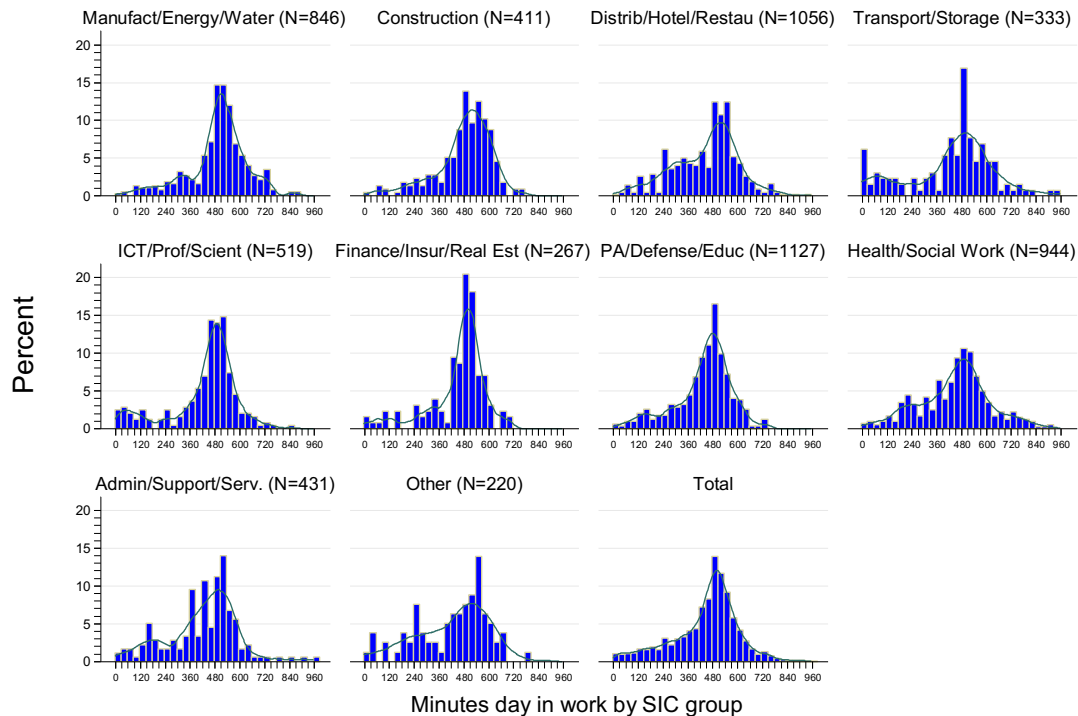
Work duration by occupation



Minutes day in work by SOC 1 digit

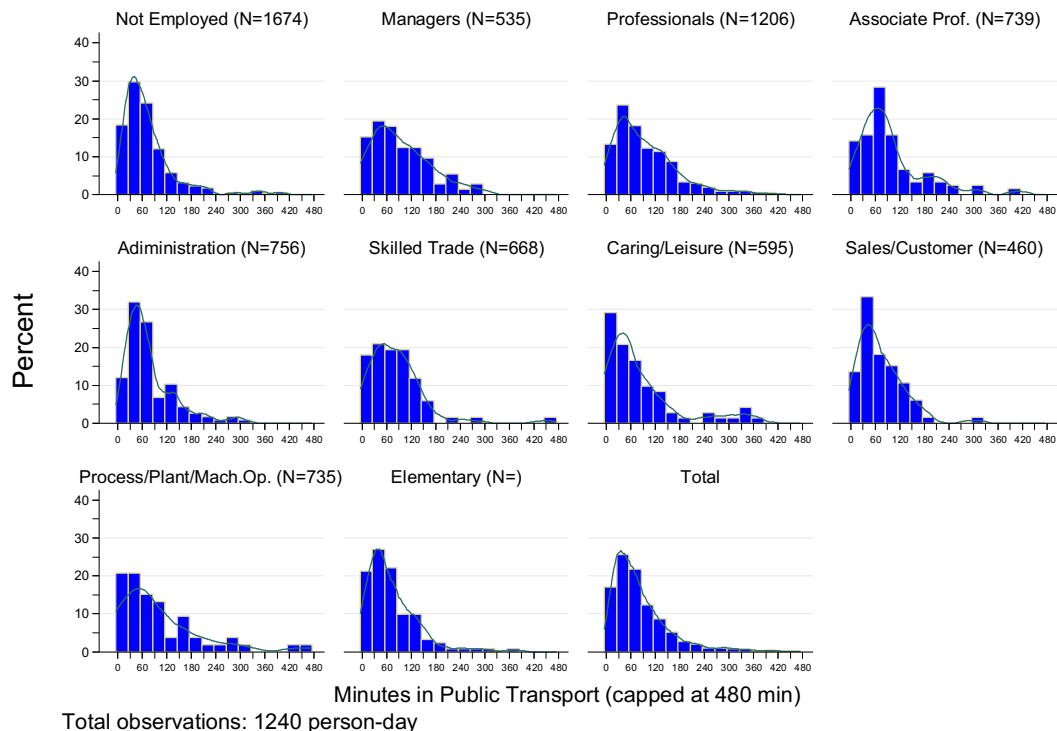
Total observations: 7828 person-day

Work duration by *industry*

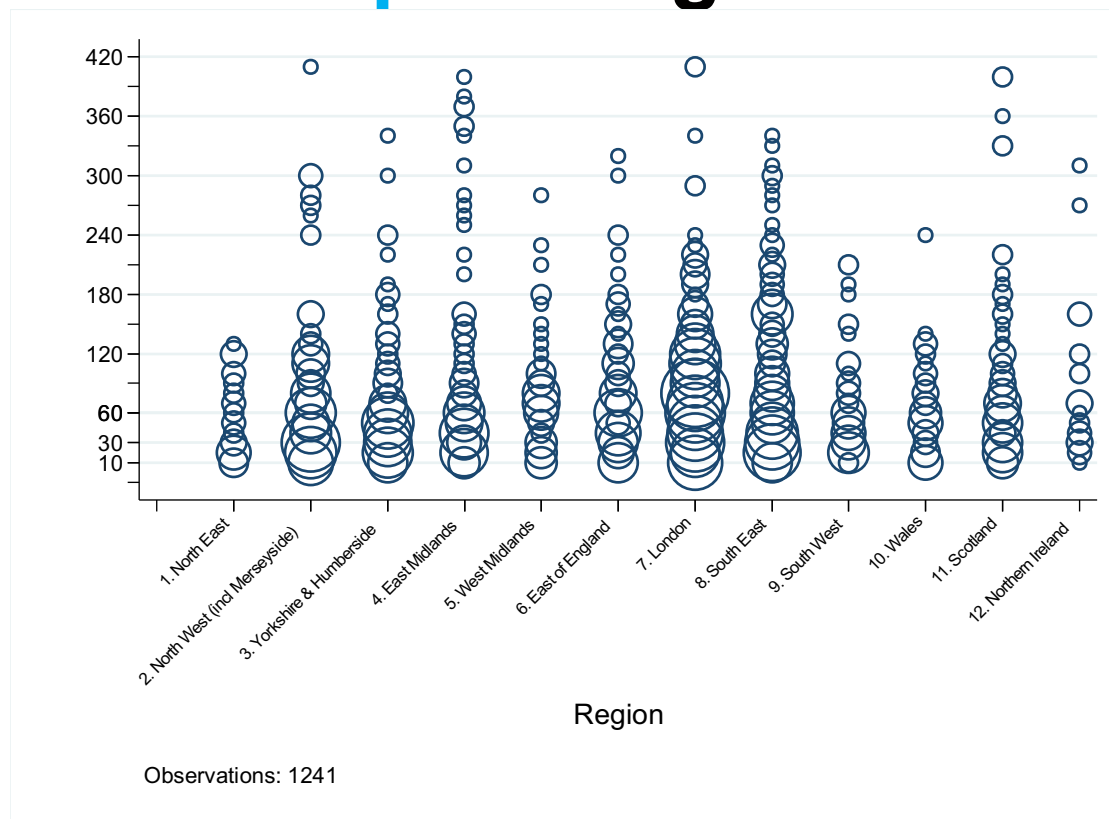


Total observations: 2664 person-day

Public transport use by occupation



Public transport: regional variation



The
Alan Turing
Institute

Nik Lomax

nlomax@turing.ac.uk



turing.ac.uk
@turinginst

Jamil Nur
jn381@cam.ac.uk



**Karyn
Morrissey**
@turing.ac.uk

