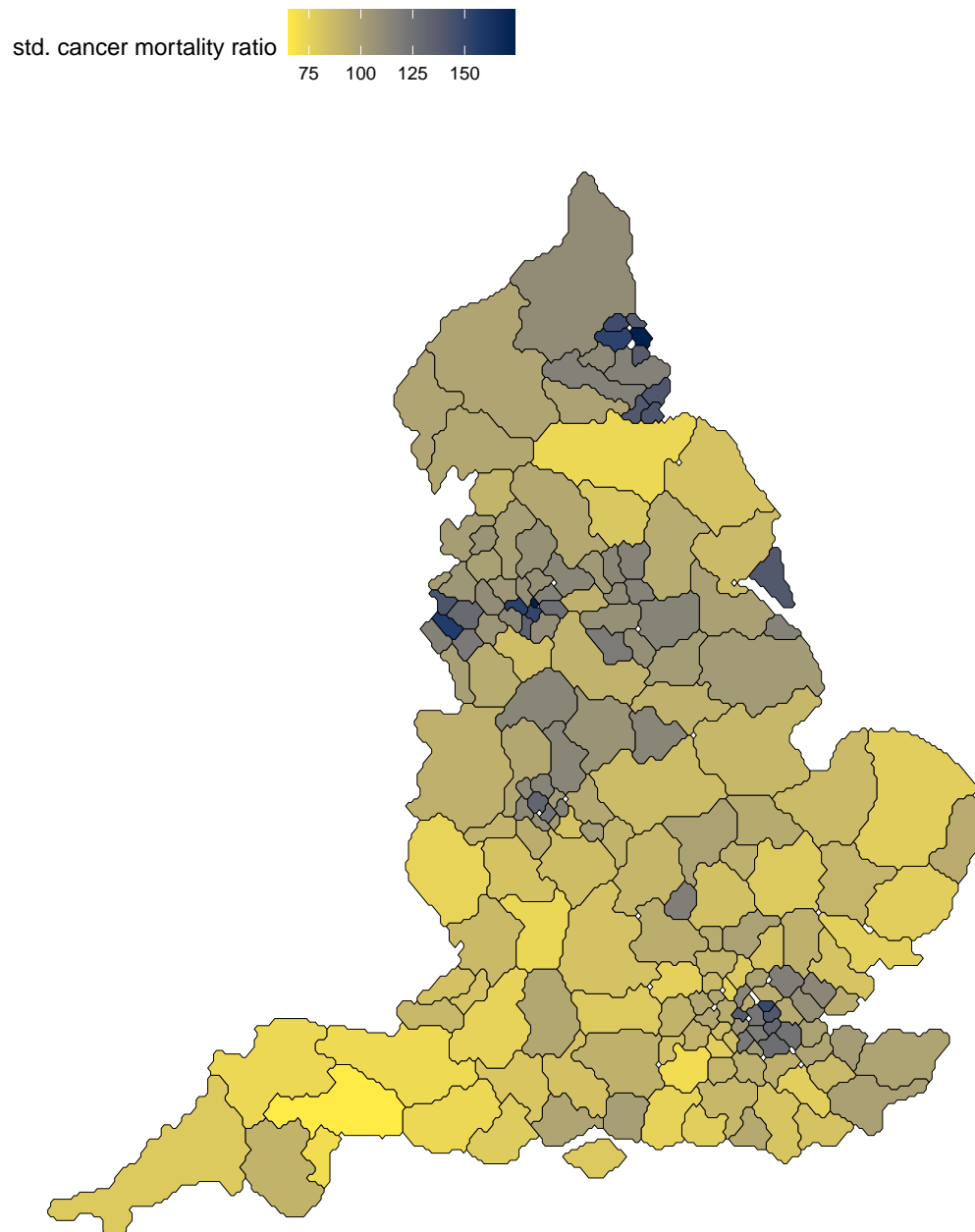


Lab 4

Antonio Jurlina

10/05/2020

Standardized Mortality Ratios for Lung Cancer (SMRLC)



OLS Model

term	estimate	std.error	statistic	p.value
(Intercept)	29.4139045	8.4700247	3.472706	0.0006408
x	0.0002972	0.0001640	1.812242	0.0715615
y	0.0006587	0.0000984	6.695058	0.0000000
jarman_sc	0.4565039	0.0576829	7.914022	0.0000000

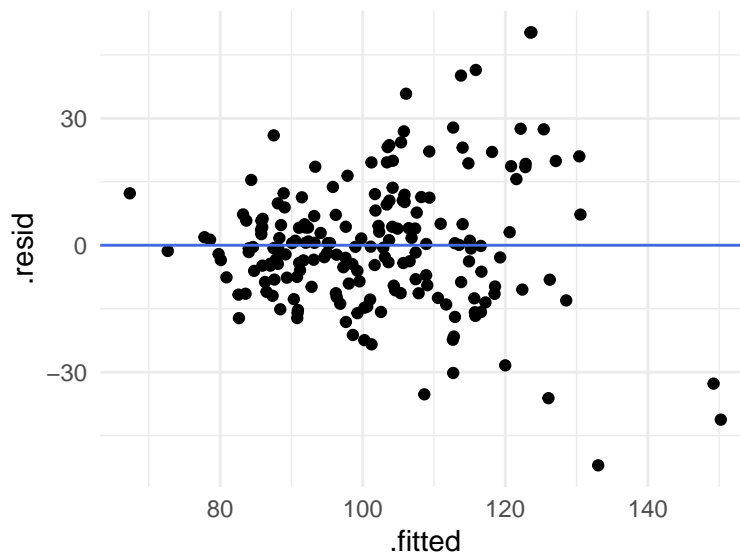
r.squared	adj.r.squared	sigma	statistic	p.value	df	logLik	AIC	BIC	deviance	df.residual	nobs
0.4509147	0.4420585	15.41211	50.91506	0	3	-787.2561	1584.512	1600.747	44181.15	186	190

The OLS model, in which

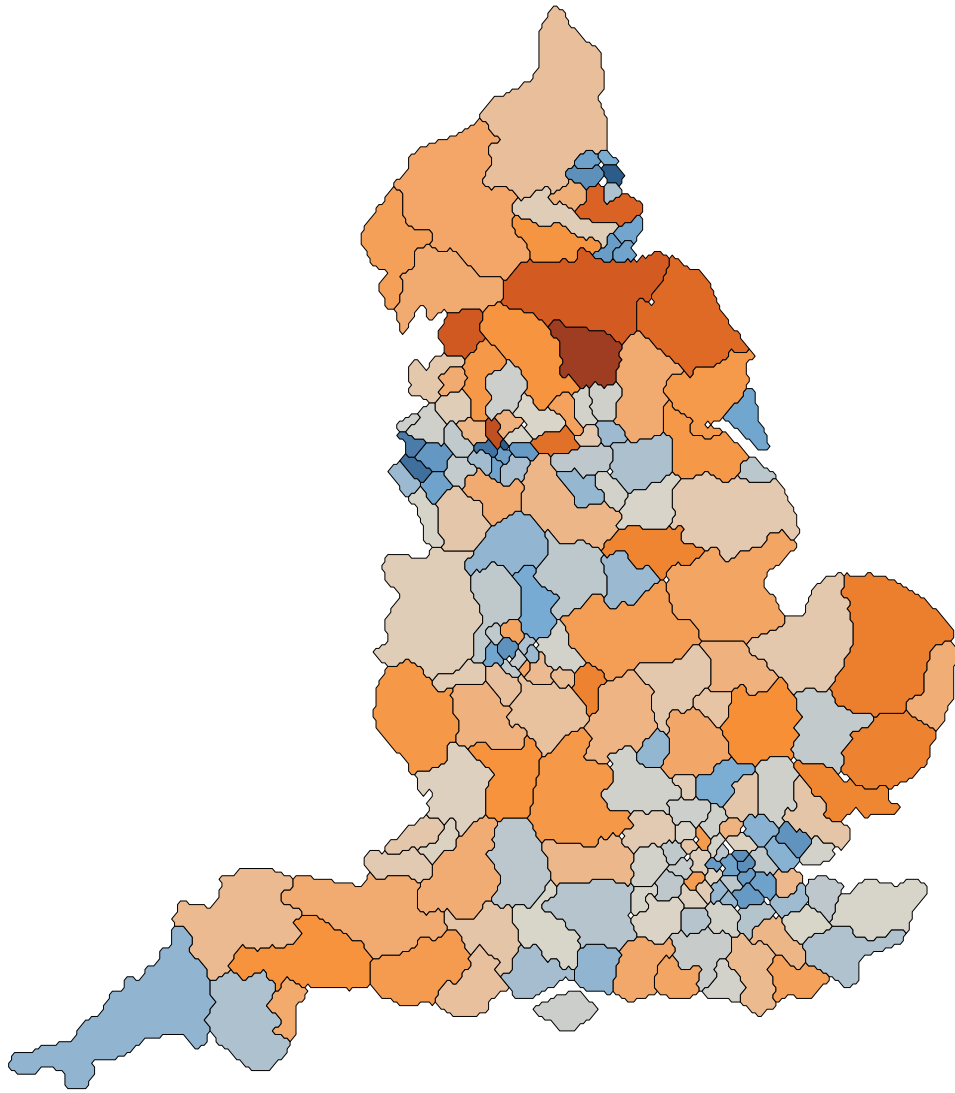
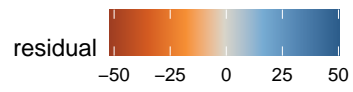
$$SMRLC = \beta_0 + \beta_1 * X + \beta_2 * Y + \beta_3 * JarmanIndex$$

shows that the spatial effects (X and Y) are not substantial, while the Jarman Index serves as a much better predictor of lung cancer mortality across UK. According to the model, each unit increase in the Index causes a 0.46 unit increase in the standardized lung cancer mortality ratio, on average.

statistic	p.value	parameter	method
45.41049	0.0000000	3	studentized Breusch-Pagan test
17.63913	0.0001478	2	Jarque Bera Test



According to the Breusch-Pagan test, the homoskedasticity of errors null hypothesis is rejected. Further, this is verifiable through the plot of residuals against the fitted values. Finally, the Jarque Bera Test also indicates that the errors are heteroskedastic. This poses a problem for the OLS approach given that the errors are not homoskedastic. It means we cannot rely on the reported values of the standard errors on model estimates, even though the estimates themselves remain unbiased.

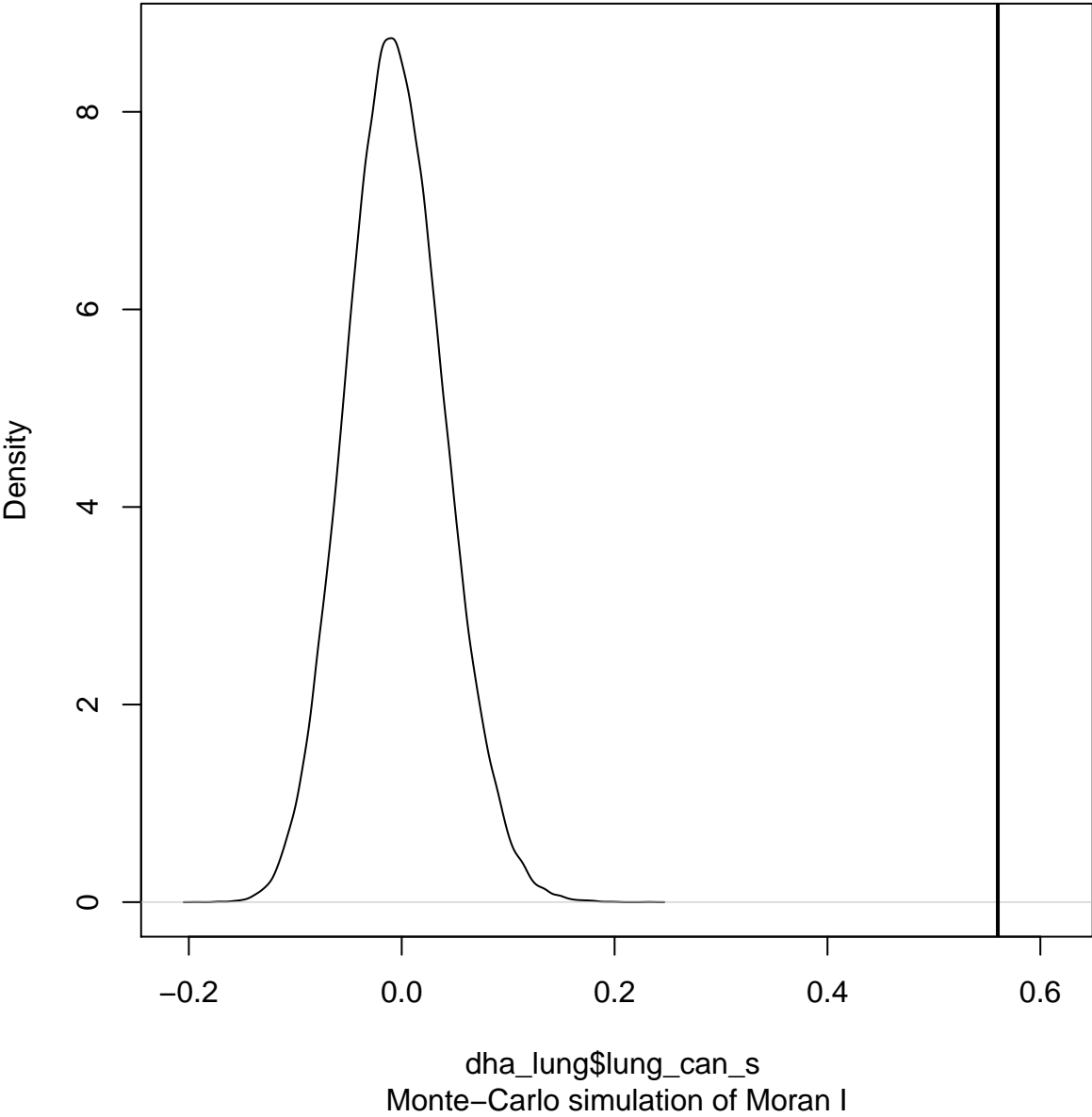


Moran’s I (spatial autocorrelation)

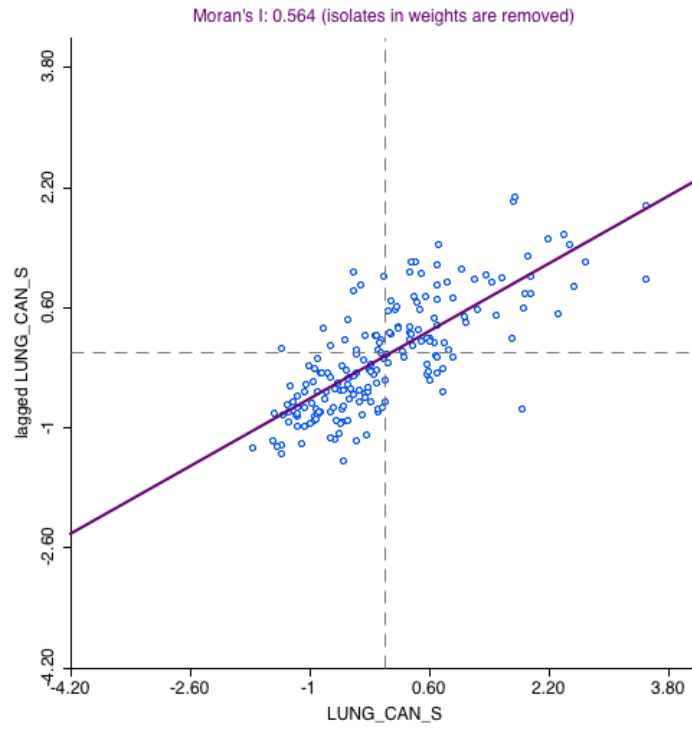
estimate1	estimate2	estimate3	statistic	p.value	method	alternative
0.5600069	-0.0053191	0.0021292	12.25154	0	Moran I test under randomisation	greater

statistic	p.value	parameter	method	alternative
0.5600069	1e-05	1e+05	Monte-Carlo simulation of Moran I	greater

Density plot of permutation outcomes



The underlying distribution of lung cancer mortality ratios is assumed to be the result of a random spatial process, in our null hypothesis. However, the statistically significant value of 0.56 from Moran's I, has us reject the null and assume that the high values are clustered together due to something other than a simple random spatial process.



```

## R version 3.6.2 (2019-12-12)
## Platform: x86_64-apple-darwin15.6.0 (64-bit)
## Running under: macOS Sierra 10.12.6
##
## Matrix products: default
## BLAS: /Library/Frameworks/R.framework/Versions/3.6/Resources/lib/libRblas.0.dylib
## LAPACK: /Library/Frameworks/R.framework/Versions/3.6/Resources/lib/libRlapack.dylib
##
## locale:
## [1] en_US.UTF-8/en_US.UTF-8/en_US.UTF-8/C/en_US.UTF-8/en_US.UTF-8
##
## attached base packages:
## [1] stats      graphics  grDevices  utils      datasets  methods    base
##
## other attached packages:
## [1] tseries_0.10-47  lmtest_0.9-37   zoo_1.8-8      ggthemes_4.2.0
## [5] rgdal_1.5-16     classInt_0.4-3  spdep_1.1-5    spData_0.3.8
## [9] rebus_0.1-3      lubridate_1.7.9 tmap_3.1       sf_0.9-5
## [13] sp_1.4-2         forcats_0.5.0   stringr_1.4.0  dplyr_1.0.2
## [17] purrr_0.3.4      readr_1.3.1     tidyr_1.1.2    tibble_3.0.3
## [21] ggplot2_3.3.2    tidyverse_1.3.0
##
## loaded via a namespace (and not attached):
## [1] leafem_0.1.3      colorspace_1.4-1  deldir_0.1-29
## [4] ellipsis_0.3.1    class_7.3-17      rprojroot_1.3-2
## [7] leaflet_2.0.3     base64enc_0.1-3   fs_1.5.0
## [10] dichromat_2.0-0   rstudioapi_0.11   farver_2.0.3
## [13] fansi_0.4.1       xml2_1.3.2        codetools_0.2-16
## [16] splines_3.6.2     knitr_1.29        jsonlite_1.7.0
## [19] tmaptools_3.1     broom_0.7.0       dbplyr_1.4.4
## [22] png_0.1-7         compiler_3.6.2    httr_1.4.2
## [25] backports_1.1.9   assertthat_0.2.1  Matrix_1.2-18
## [28] cli_2.0.2         htmltools_0.5.0   tools_3.6.2
## [31] coda_0.19-3       gtable_0.3.0      glue_1.4.2
## [34] rebus.base_0.0-3  gmodels_2.18.1    Rcpp_1.0.5
## [37] cellranger_1.1.0  raster_3.3-13     vctrs_0.3.4
## [40] gdata_2.18.0      nlme_3.1-149      leafsync_0.1.0
## [43] crosstalk_1.1.0.1 lwgeom_0.2-5      xfun_0.16
## [46] rebus.datetimes_0.0-1 rvest_0.3.6       lifecycle_0.2.0
## [49] rebus.numbers_0.0-1 gtools_3.8.2      XML_3.99-0.3
## [52] LearnBayes_2.15.1 MASS_7.3-52        scales_1.1.1
## [55] hms_0.5.3         parallel_3.6.2    expm_0.999-5
## [58] RColorBrewer_1.1-2 curl_4.3           quantmod_0.4.17
## [61] yaml_2.2.1        stringi_1.4.6     highr_0.8
## [64] e1071_1.7-3       TTR_0.24.2        boot_1.3-25
## [67] rlang_0.4.7       pkgconfig_2.0.3   evaluate_0.14
## [70] lattice_0.20-41   labeling_0.3       htmlwidgets_1.5.1
## [73] tidyselect_1.1.0  here_0.1           magrittr_1.5
## [76] R6_2.4.1          generics_0.0.2    DBI_1.1.0
## [79] pillar_1.4.6      haven_2.3.1        withr_2.2.0
## [82] xts_0.12-0        units_0.6-7        stars_0.4-3
## [85] abind_1.4-5       rebus.unicode_0.0-2 modelr_0.1.8
## [88] crayon_1.3.4      KernSmooth_2.23-17 rmarkdown_2.3
## [91] grid_3.6.2        readxl_1.3.1      blob_1.2.1

```

```
## [94] reprex_0.3.0      digest_0.6.25      munsell_0.5.0
## [97] viridisLite_0.3.0  quadprog_1.5-8
```