

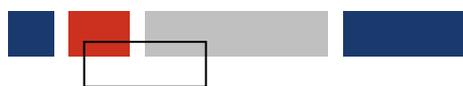


INSTITUTO NACIONAL DE ESTATÍSTICA
STATISTICS PORTUGAL

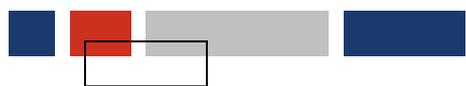
» Mismatch between jobs and skills in the EU

JOCLAD 2017 «

DMSI-ME / Joao S. Lopes

 (21-04-2017)

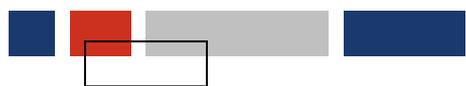




Motivation



- “Skills development are essential in the emerging new economy and fast-changing labour market”¹
- “Qualification and skill mismatches entail significant economic and social costs for individuals and firms”¹
- Skills mismatch (i.e. over-qualification, under-qualification) remains at 45% (CEDEFOP, 2015)²
- EU Guidelines (2015) call for enhancing labour supply, skills and competences³



Motivation



Create framework that:

1. combines **Official Statistics** with **Big Data**
2. estimates **Labour Market Attractiveness** and its association with **Skills Mismatch**, Labour Market Mobility and Emigration
3. is aimed at **policy makers** and both **jobseekers** and **job providers**

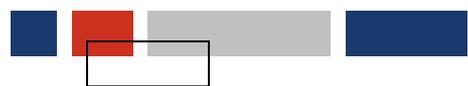


Data: LMkt Attractiveness



- “reg_dem” – demographic statistics
- “earn” – earning structure
- “educ_uoe_fin” – public expenditure on education
- “ilc” – income and life conditions
- “employ” – employment statistics
- “nama10” – annual national accounts
- “educ_part” – participation in education

7 datasets, 17 main variables



Data: LMkt Attractiveness



- “reg_dem” by **age (NUTS2)**
- “earn” by **occupation** and **economic activity**
- “educ_uoe_fin”
- “ilc” (**NUTS2**)
- “employ” by **age, education level, economic activity (NUTS2)**
- “nama10” (**NUTS2**)
- “educ_part” (**NUTS2**)

7 datasets, **17** main variables, **76** variables

subjects: **NUTS0 = 28**; NUTS1= 98; NUTS2 = 276.





Data: Skills mismatch



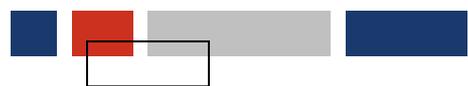
- “EURES” **scrapped data** on jobseekers’ CVs
- “EURES” **scrapped data** on job vacations

2 datasets, 1 main variable

... but

cleaning and **structuring** requires considerable expertise

normalization requires detailed demographical information

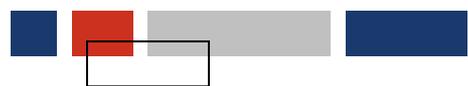


Data: Skills mismatch



- “educ_uoe_grad02” – education statistics
- “jvs_a_nace2” – job vacancy statistics

2 datasets, 1 main variable



Data: Skills mismatch



- “educ_uoe_grad02” by **education field**
- “jvs_a_nace2” by **occupation** and **economic activity (NUTS2)**

2 datasets, 1 main variable, 1 variable

subjects: **NUTS0 = 8**; NUTS1= 14; NUTS2 = 47



Methods



- Network Analysis
- Partition-around-medoids (PAM)⁴
- Over-representation analysis (ORA)
- Multinomial regression
- Multivariate regression
- Weighted correlation network analysis (WCNA)⁵



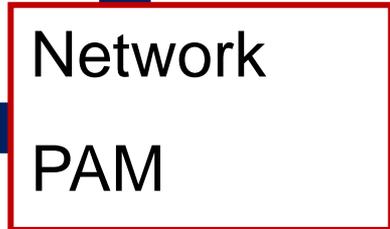
Methods



Labour Market Attractiveness (by NUTS 0-2)



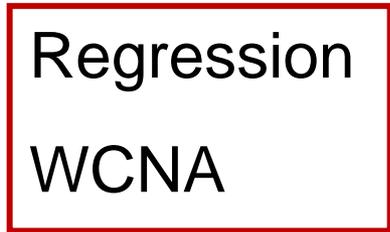
distance between NUTS



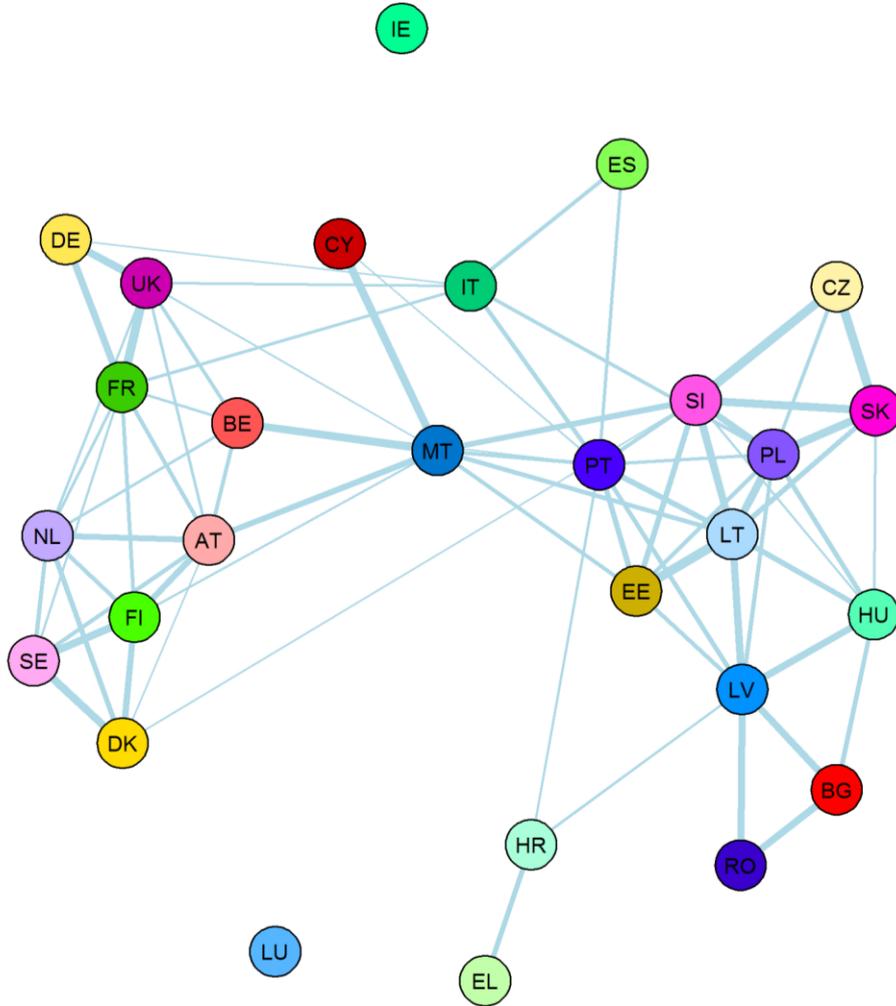
NUTS groups



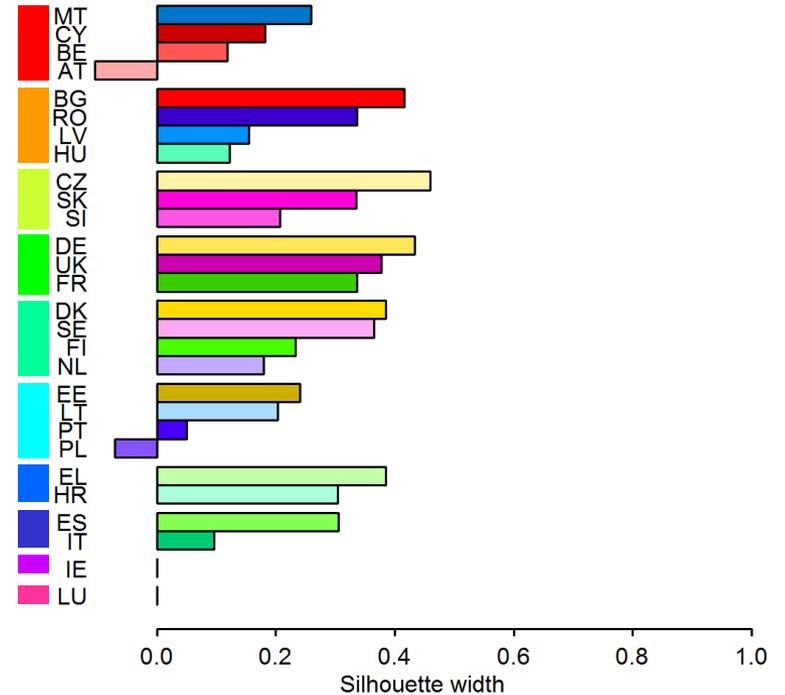
Skills mismatch (by NUTS0-2)



Results: LMkt Attractiveness



Labour market attractiveness





Results: LMkt Attractiveness



“MT-CY-BE-AT”

emp_15-24_ED5-8
 expend_ED5-8

“BG-RO-LV-HU”

ARPR_socexcl
 emp_Y15-24_NaceA
 earn_OC[1-9]

“CZ-SK-SI”

emp_FT
 pop_Y25-64
 ARPR_socexcl

“DE-UK-FR”

emp_PT
 disp_income
 pop_Total

“DK-SE-FI-NL”

earn_OC[1-9]

expend_ED5-8

ARPR_socexcl

“EE-IT-PT-PL”

earn_OC[1-5,7-9]

“EL-HR”

GVAgr

“ES-IT”

emp_Y15-24_Nace[A,J-L]
 expend_ED5-8

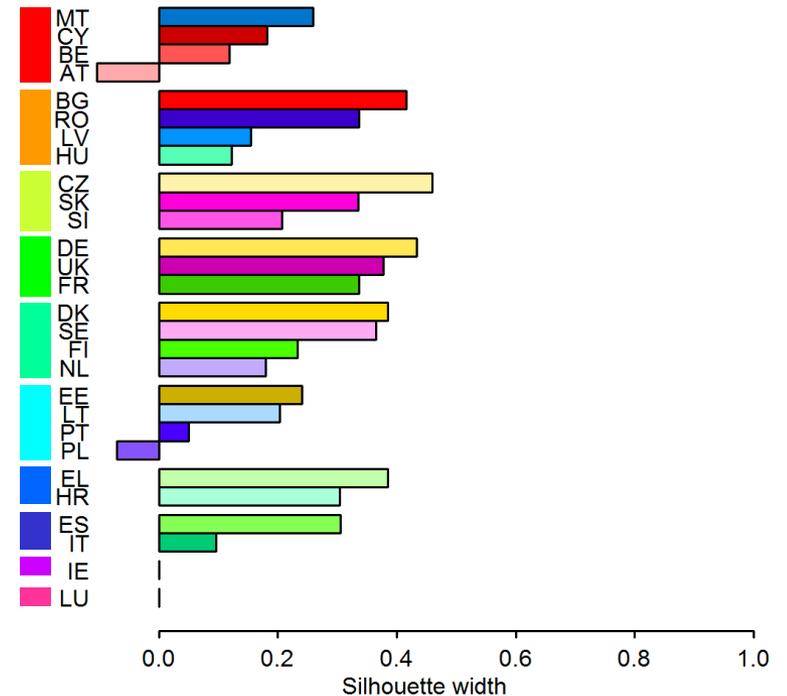
“IE”

GDP
 GVAgr
 pop_Y0-14

“LU”

emp_Y25-64_ED5-8
 GDP
 low_work

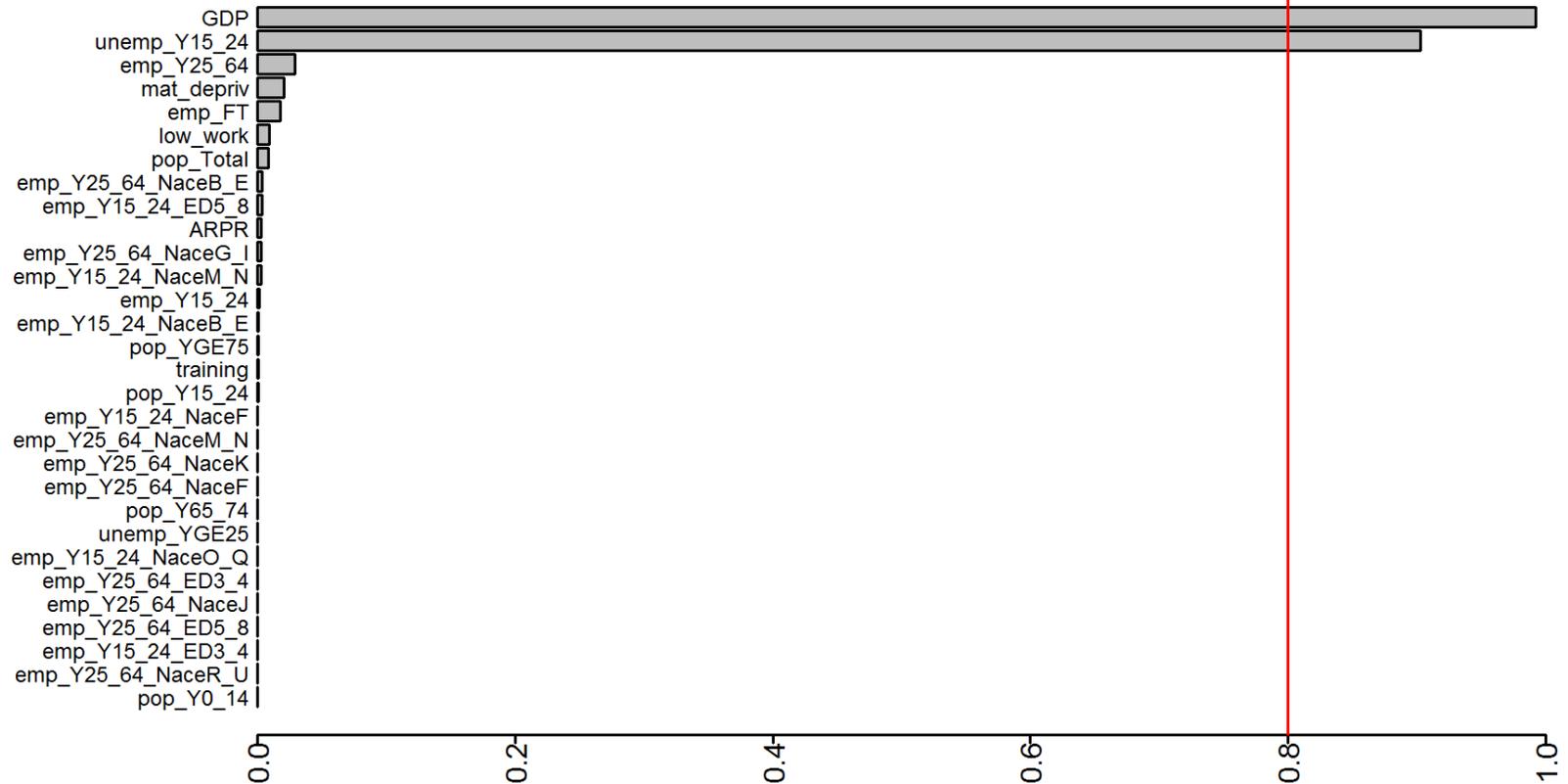
Labour market attractiveness



Results: LMkt Attractiveness



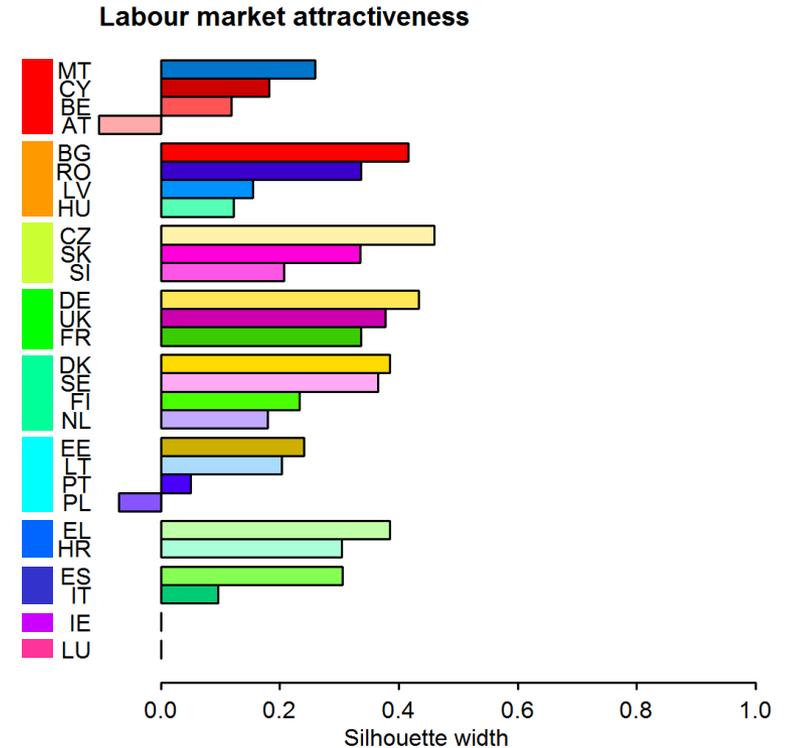
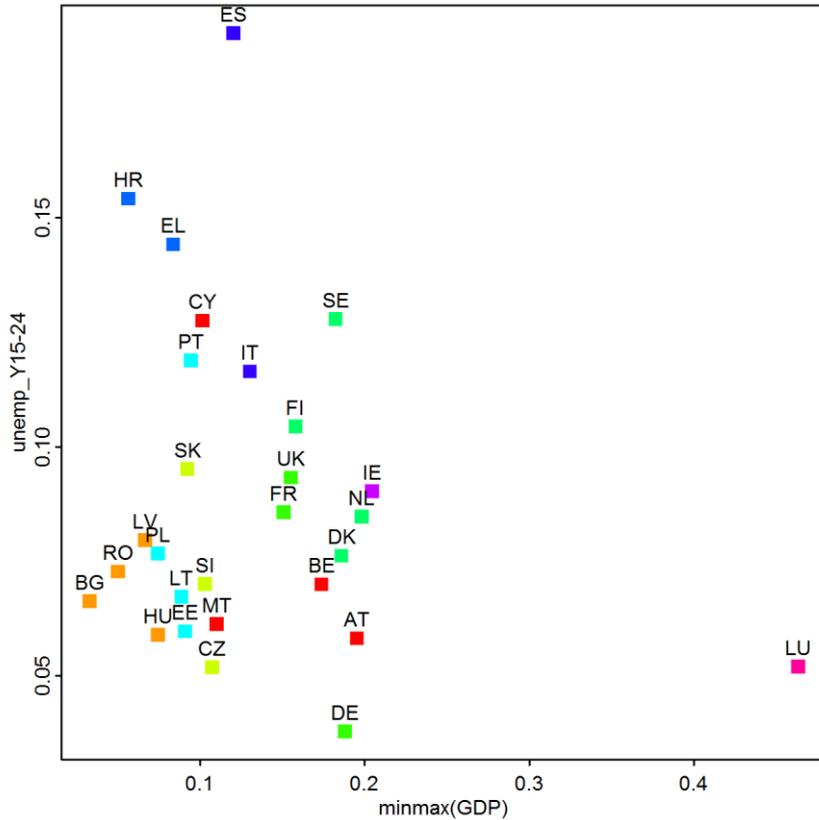
Model-averaged importance of terms



EU_groups ~ 1 + unemp_Y15-24 + GDP

R_{McFadden} = 0.82, R_{count} = 0.82

Results: LMkt Attractiveness



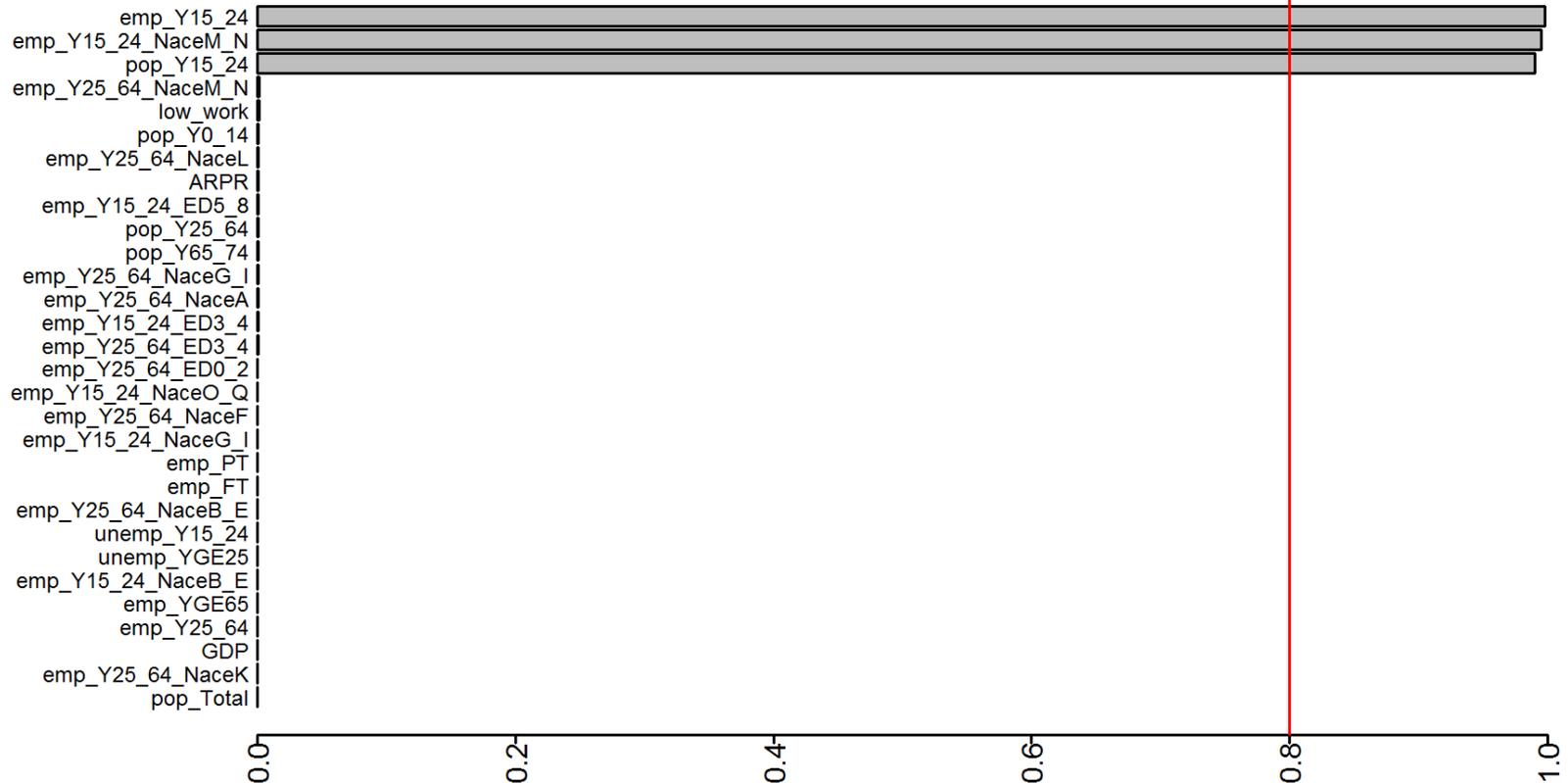
EU_groups ~ 1 + unemp_Y15-24 + GDP

R_{McFadden} = 0.82, R_{count} = 0.82

Results: Skills mismatch



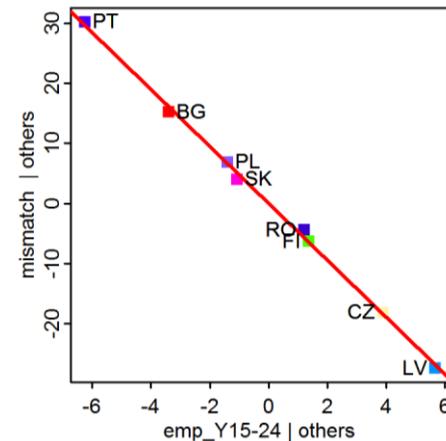
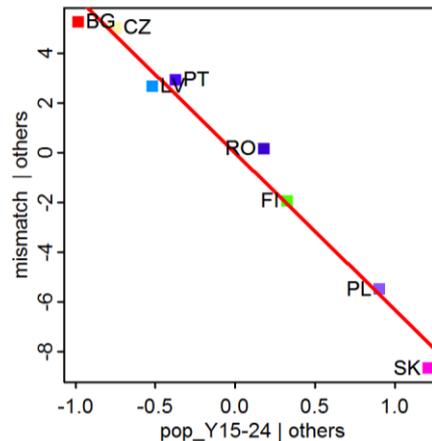
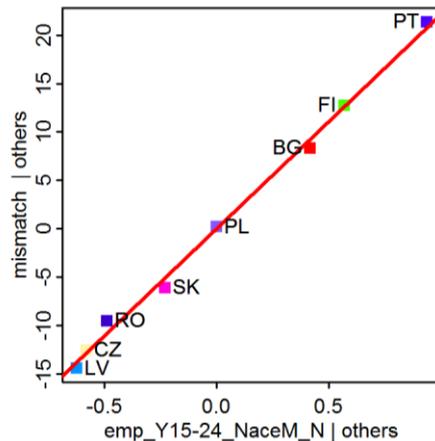
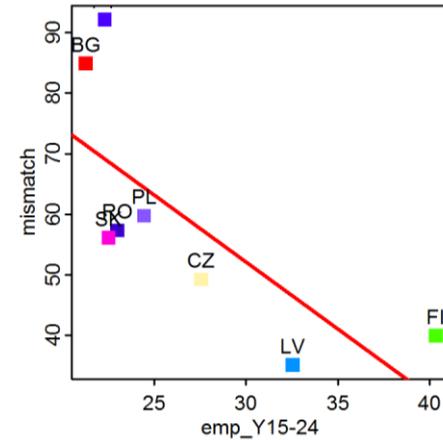
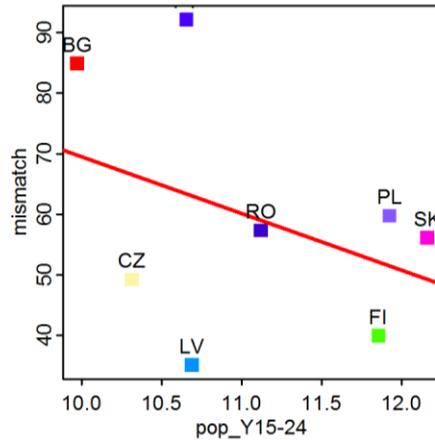
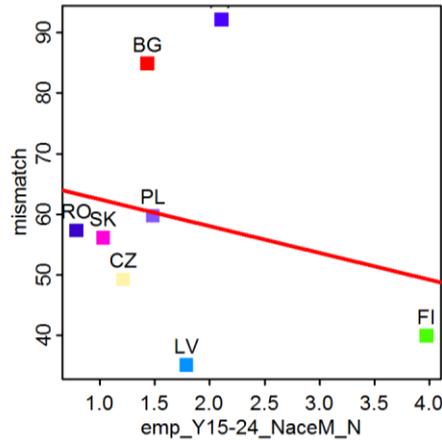
Model-averaged importance of terms



$$\text{mismatch} = 217.8 + 22.2 * \text{emp_Y15-24_NaceM_N} - (6.3 * \text{pop_Y15-24} + 4.7 * \text{emp_Y15-24})$$

$$R^2 = 1.00$$

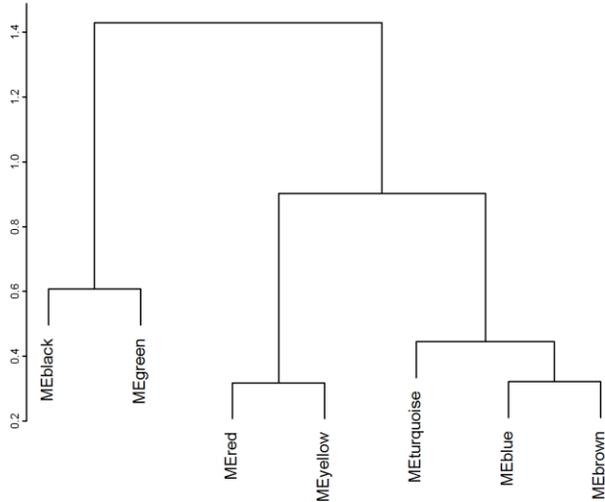
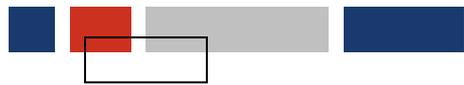
Results: Skills mismatch



$$\text{mismatch} = 217.8 + 22.2 \cdot \text{emp_Y15-24_NaceM_N} - (6.3 \cdot \text{pop_Y15-24} + 4.7 \cdot \text{emp_Y15-24})$$

$R^2 = 1.00$

Results: WCNA



MEblack

unemp_Y[15-24,25-64]

MEgreen

ilc_ARPR

ilc_ARPR_socexcl

pop_YG75

MEred

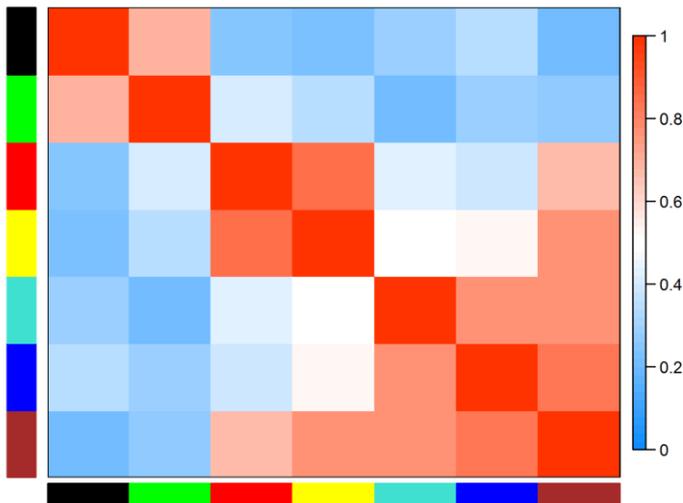
emp_Y[15-24,25-64]_NaceB-E

MEyellow

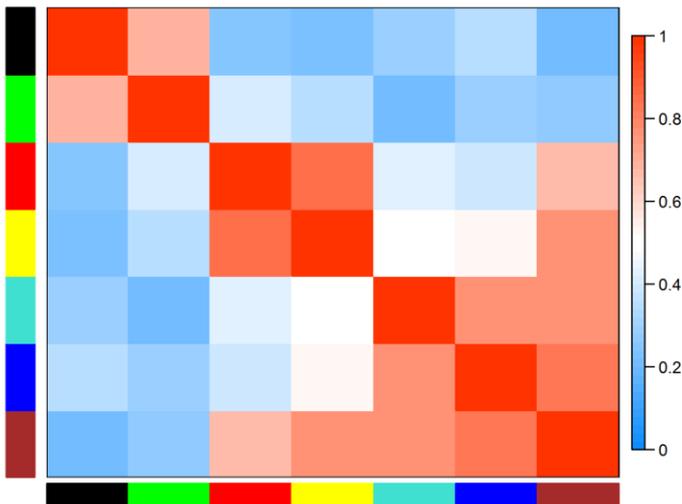
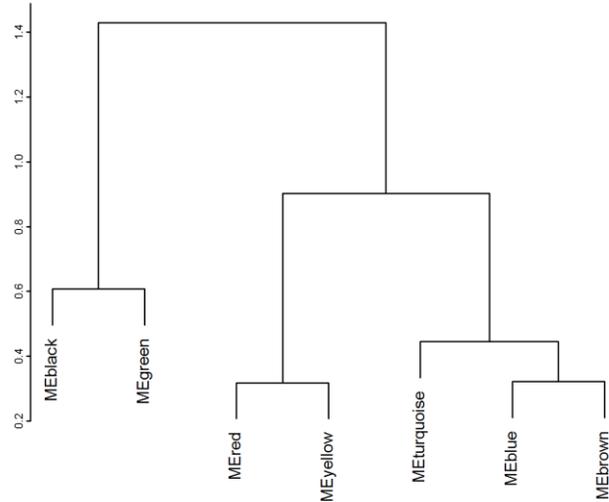
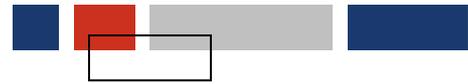
ilc_low_work

emp_Y25-64_ED[0-2,3-4]

emp_Y25-64_NaceF



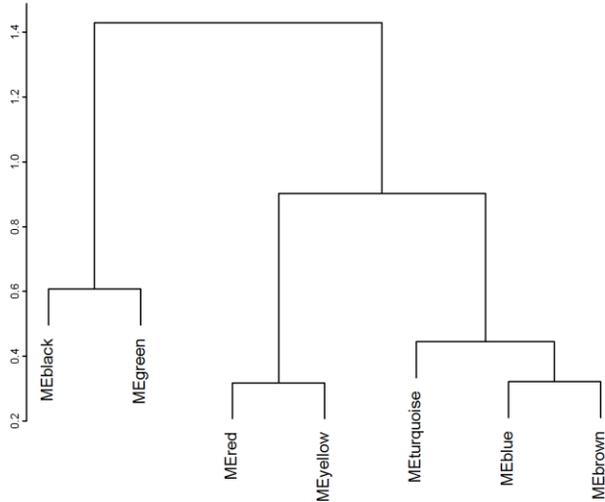
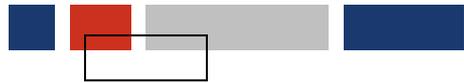
Results: WCNA



MEturquoise

- ilc_mat_depriv
- ilc_rooms_pp
- earn_OC[1-9]
- emp_[FT,PT]
- emp_Y15-24_ED0-2
- emp_Y15-24_Nace[A,O-Q,R-U]
- emp_Y25-64_ED5-8
- emp_Y25-64_Nace[A,G-I,J,K,M_N,O-Q,R-U]
- na_disp_income
- na_GDP
- pop_Y[0-14,25-64]
- training

Results: WCNA

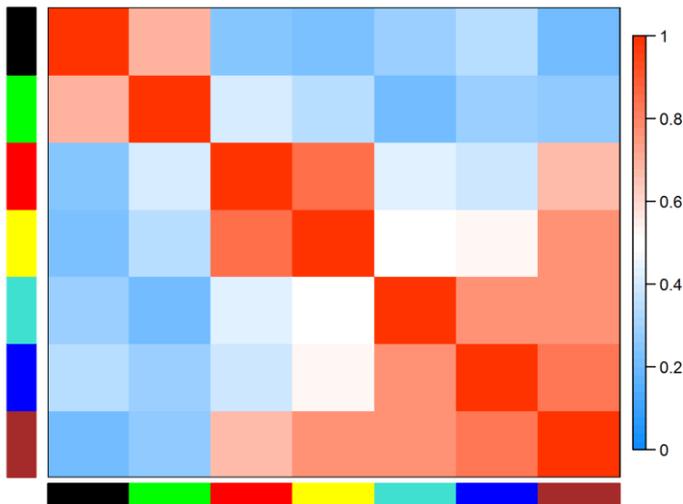


MEblue

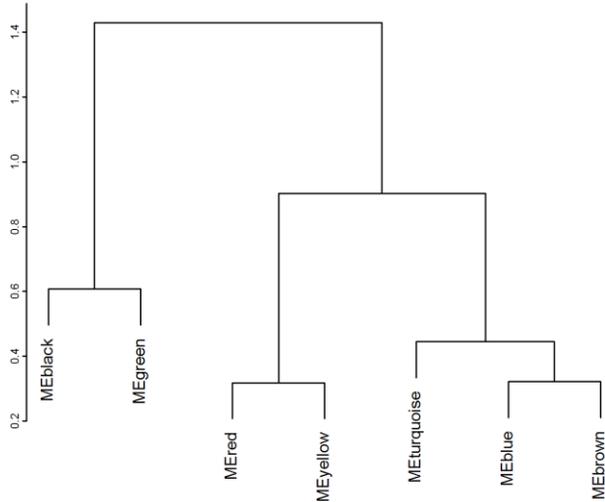
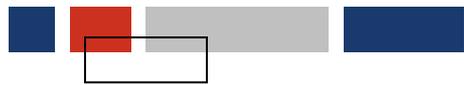
emp_Y15-24_ED5-8
 emp_Y15-24_Nace[G-I,J,M_N]
 emp_YGE65
 pop_Y15-24

MEbrown

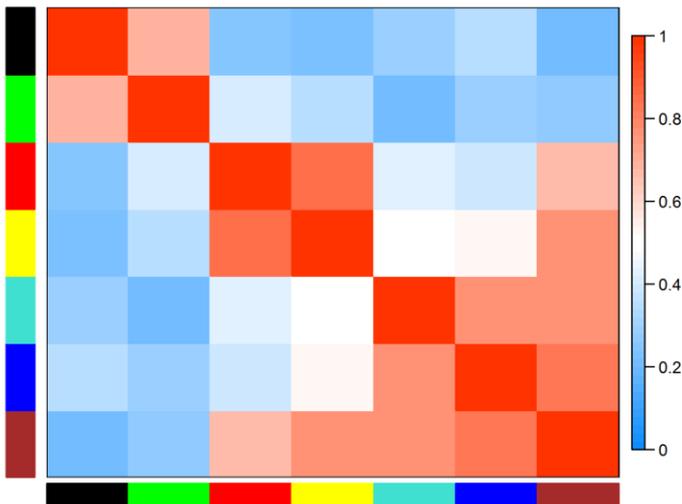
emp_Y[15-24,25-64]
emp_Y15-24_ED3-4
 emp_Y15-24_NaceF
 emp_Y25-64_NaceL

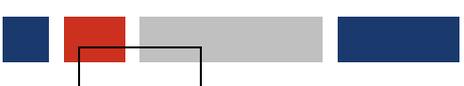


Results: WCNA



labels	description	mismatch	EU_groups
MEblack	Unemployment		
MEgreen	Poverty		0.63
MEred	Industry sector		
MEyellow	emp_Y25-64_ED0-4	-0.50	
MEturquoise	Earnings		0.80
MEblue	emp_Y15-24_ED5-8		0.63
MEbrown	emp_Y15-24_ED3-4	-0.86	0.57





Conclusions

- 
- **LMkt Attract** is able to form consistent clusters at NUTS0
 - **Youth unemployment** and **GDP** can separate well clusters
 - **LMkt Attract** can be separate in different modules: **Unemployment**, **Poverty**, **Industry**, **emp_Y25-64_ED0-4**, **Earnings**, **emp_Y15-24_ED3-4** and **emp_Y15-24_ED5-8**
 - **Skills Mismatch** is associated with **population Y15-24**, negative association with **Pop_{prop}**, **Emp_{prop}** and **NaceM_N_{prop}**
 - **Skills Mismatch** is strongly associated to module **emp_15-24_ED3-4**



Acknowledge



Team:

Sónia Quaresma

João Lopes

Marco Moura

Institution:

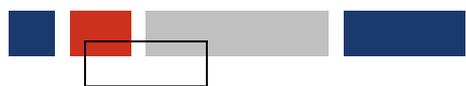


Data:



Thank you!





Bibliography



1. https://ec.europa.eu/commission/publications/skills-education-and-lifelong-learning-european-pillar-social-rights_en
2. CEDEFOP (2015) “Skills, qualifications and jobs in the EU: the making of a perfect match? “
3. Council Decision (EU) 2015/1848 of 5 October 2015
4. Reynolds et al. (1992) “Clustering rules: A comparison of partitioning and hierarchical clustering algorithms” J Math. Model. Algorithms
5. Langfelder and Horvath (2008) “WGCNA: an R package for weighted correlation network analysis” BMC Bioinformatics



Tools



R libraries

car - Companion to Applied Regression

caret - Classification and Regression Training

cluster - Finding Groups in Data: Cluster Analysis Extended

glmulti - Model selection and multimodel inference made easy

MASS - Support Functions and Datasets for MASS

nnet - Feed-Forward Neural Networks and Multinomial Log-Linear Models

sna - Tools for Social Network Analysis

WGCNA - Weighted Correlation Network Analysis