

Regional Differences and Small-area Estimation of Mortality in Taiwan

Jack Yue

National Chengchi University

Date: Sept. 16, 2024

Email: csyue@nccu.edu.tw



Regional Disparity in Taiwan's Lifespan

- Taiwan's life expectancy increases steadily.
 - County difference is close to 9 years (2023 male, 80.58 vs. 71.91), reducing slightly.
 - Mortality differences in major causes of death are also significant between counties and township.

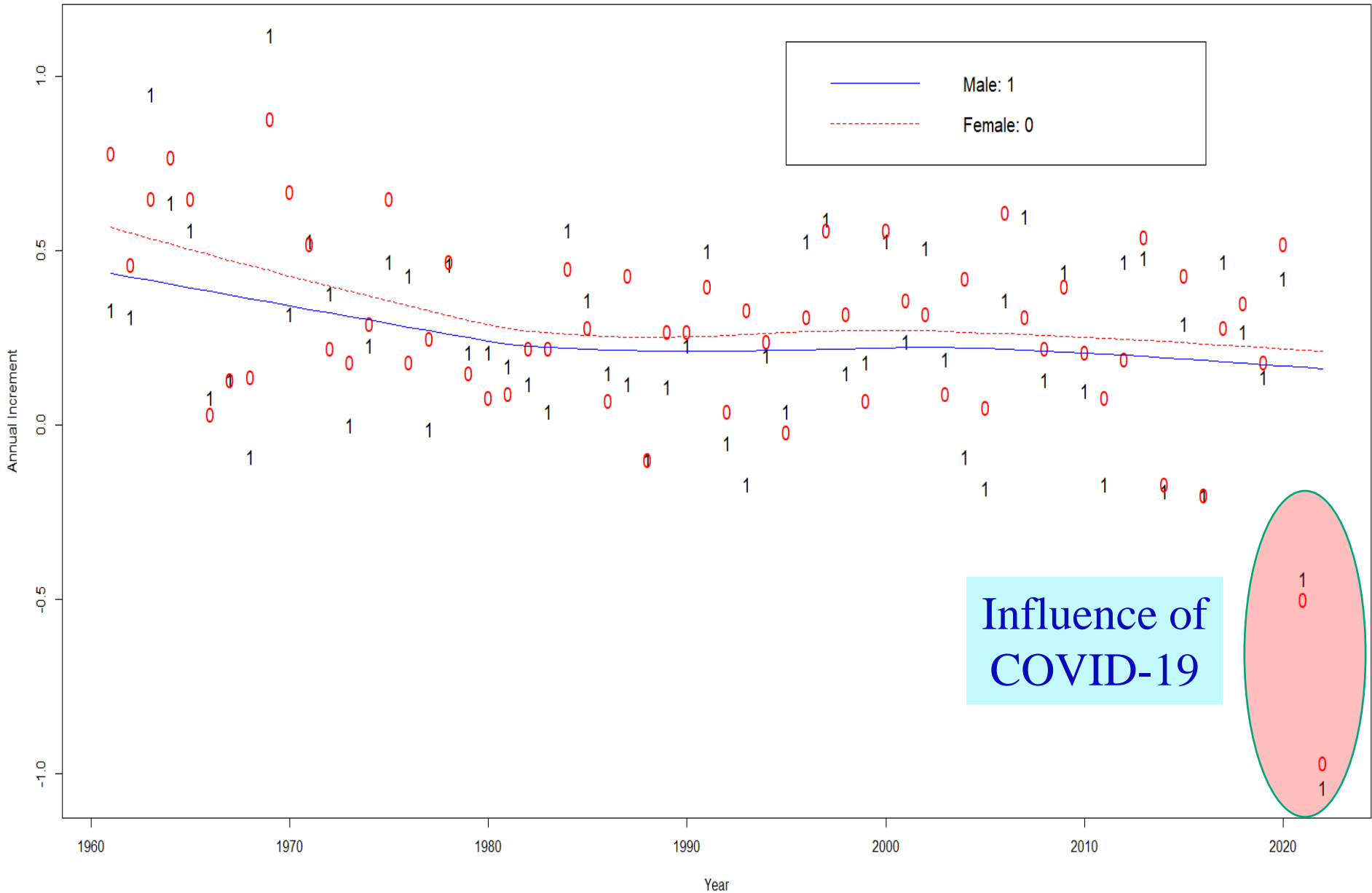


<https://mobilephysiotherapyclinic.in/cancer/>



<https://www.fideliscare.org/Member/Helpful-Tools/Health-Resources/Heart-Disease>

Annual Increment of Life Expectancy in Taiwan



Life Expectancy of Taiwan's Complete Life Tables

Life Expectancy	Male		Female	
	2019-21	2009-11	2019-21	2009-11
Complete Life Table	77.41	75.98	83.83	82.33
Death Cause Eliminated				
Cancer	4.09	4.16	3.12	2.88
Heart Disease	1.41	1.37	1.12	1.11
Accident	0.85	1.08	0.41	0.48
Pneumonia	0.83	0.70	0.67	0.63
Cerebrovascular Disease	0.78	0.88	0.64	0.73
Diabetes	0.60	0.65	0.67	0.66
Chronic Liver Disease	0.43	0.62	0.16	0.26
Chronic Lower Respiratory Disease	0.41	0.52	0.20	0.23
Hypertensive Disease	0.37	---	0.39	---
Nephritis, Nephrotic Syndrome, and Nephrosis	0.27	0.33	0.32	0.34

Mortality Models for Small Areas

- Modelling small area mortality rates is necessary due to regional disparity.
- We use mortality models (e.g., Lee-Carter model) but they have large estimation errors when the population size is small.
- This happens to different estimation methods, such as SVD, MLE, or approximation method.

Lee-Carter Model

- Lee and Carter (1992) proposed a mortality model for forecasting (central) mortality rates

$$\ln(m_{xt}) = \alpha_x + \beta_x \kappa_t + \varepsilon_{xt}$$

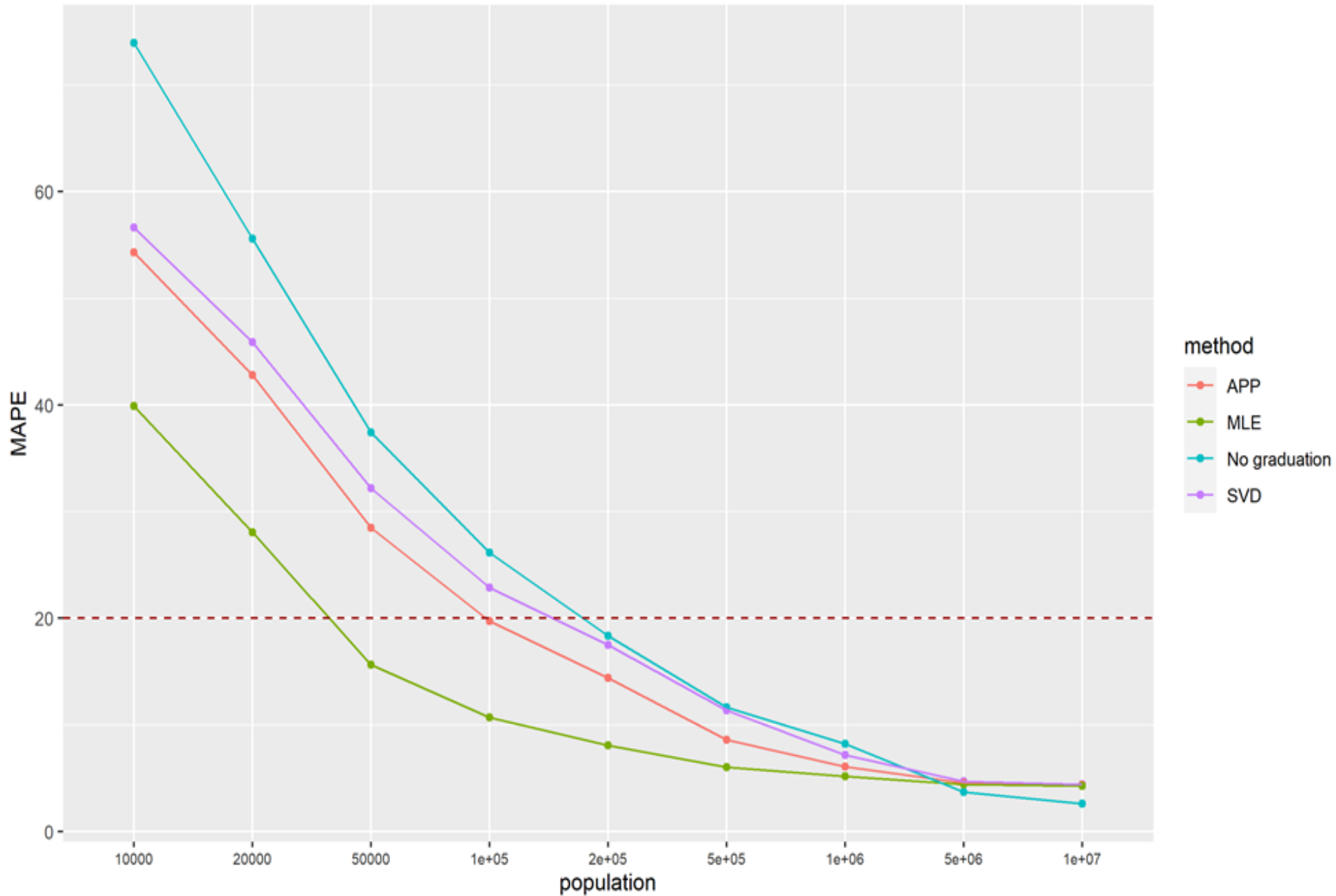
where $\kappa_t \rightarrow$ mortality index in the year t

$\alpha_x \rightarrow$ average age-specific mortality.

$\beta_x \rightarrow$ deviation in changes of mortality index

Note: $\beta_x \times \kappa_t$ can be used to measure annual change.

Estimation Error vs. Population size (LC Model)



Modifying LC Models for Small Areas

- A lot of modifications to the LC models have been proposed.
 - Increase the population size via referencing a larger population (Graduation).
 - Choose the reference population with similar mortality profile as the target (small) population.
- Q: Similarity level vs. Estimation error?

Partial SMR and LC Model

- Apply partial SMR (Wang et al, 2018) to mortality rates of small areas, mortality graduation.
 - Plug graduated mortality rates to LC model.
 - Better improvement for smaller populations.

MAPE of simulation (Taiwanese Male, 1990–2009, 5-age) (Unit: %).

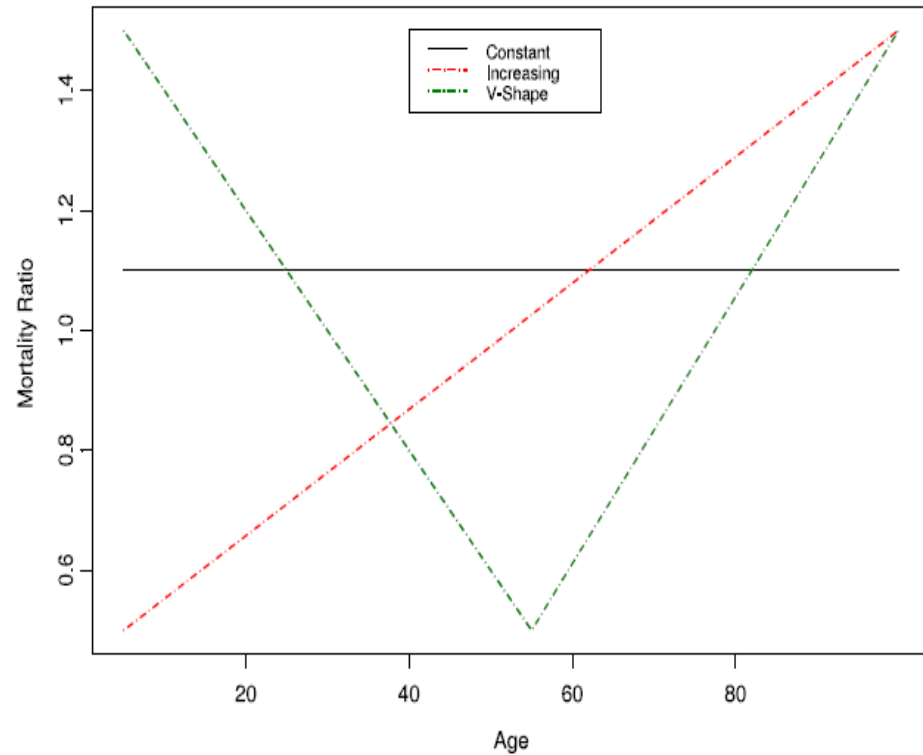
Ref. Aggregate 20-year data

	10,000	20,000	50,000	100,000	200,000	500,000
Raw	68.23	50.59	32.90	22.88	16.28	10.27
Lee-Carter	33.57	23.67	15.53	10.97	8.66	6.05
Partial SMR	14.31	11.75	9.68	8.70	8.09	7.50

Similarity Level

■ Better results for similar mortality profile.

→ Look at $\frac{m_x^S}{m_x^R}$, $\frac{\alpha_x^S}{\alpha_x^R}$, or $\frac{\beta_x^S}{\beta_x^R}$.

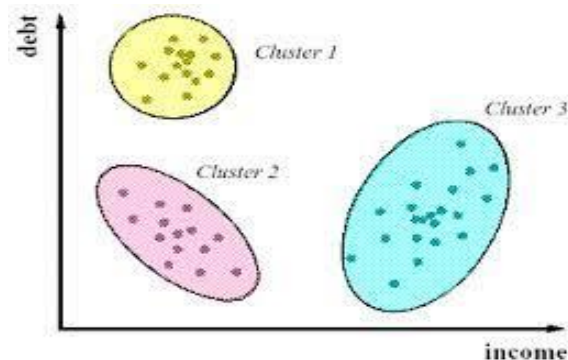


MAPE (%) of 3 mortality scenarios (Population size = 50,000).

a		0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
Constant	Raw	30.1	28.7	27.6	26.5	25.3	24.7	23.3	23.0	22.4	21.8
	Whittaker	13.2	12.8	12.5	12.3	11.9	11.7	11.5	11.4	11.1	10.9
	PSMR	5.0	4.6	4.6	4.2	4.1	4.1	3.9	3.8	3.7	3.7
Increasing	Raw	30.2	32.4	36.0	41.3	47.6	56.2	66.2	79.8	99.2	143.3
	Whittaker	13.3	15.1	18.5	23.2	28.9	36.3	45.3	57.9	78.0	122.3
	PSMR	4.9	7.5	12.7	19.4	27.4	37.0	48.8	64.9	88.8	140.9
V-shape	Raw	30.2	31.0	33.6	38.0	43.7	51.4	60.1	72.7	90.9	130.6
	Whittaker	13.3	14.2	17.1	21.5	27.1	33.7	41.4	52.4	68.7	105.8
	PSMR	4.9	7.5	12.5	18.8	26.4	35.2	45.4	59.1	79.2	123.0

Study Objective

- Look for (small) areas with similar mortality attributes as the target population, instead a single large reference population.
- Explore the mortality profile of Taiwan township, all-cause and leading death causes.
- Use Cluster Analysis to group townships into large reference populations.



Data and Methodology

- We use Taiwan's official statistics to evaluate the proposed approach.
- Death rates of all-cause and 3 leading death causes for 358 Taiwan's townships in 1992-2023.
- Using Standardized Death Rate (SDR) and Standardized Mortality Ratio (SMR) to explore the attribute of mortality rates.

EDA and CDA

- Data Analysis can be separated into two parts:
 - Exploratory Data Analysis (EDA) and Confirmatory Data Analysis (CDA)
 - EDA (or data driven) is to **learn** from the data:
Indices & Graphical tools:
 - CDA (or formal modeling) is to **confirm** the observed findings: Lee-Carter model

EDA & Data-driven

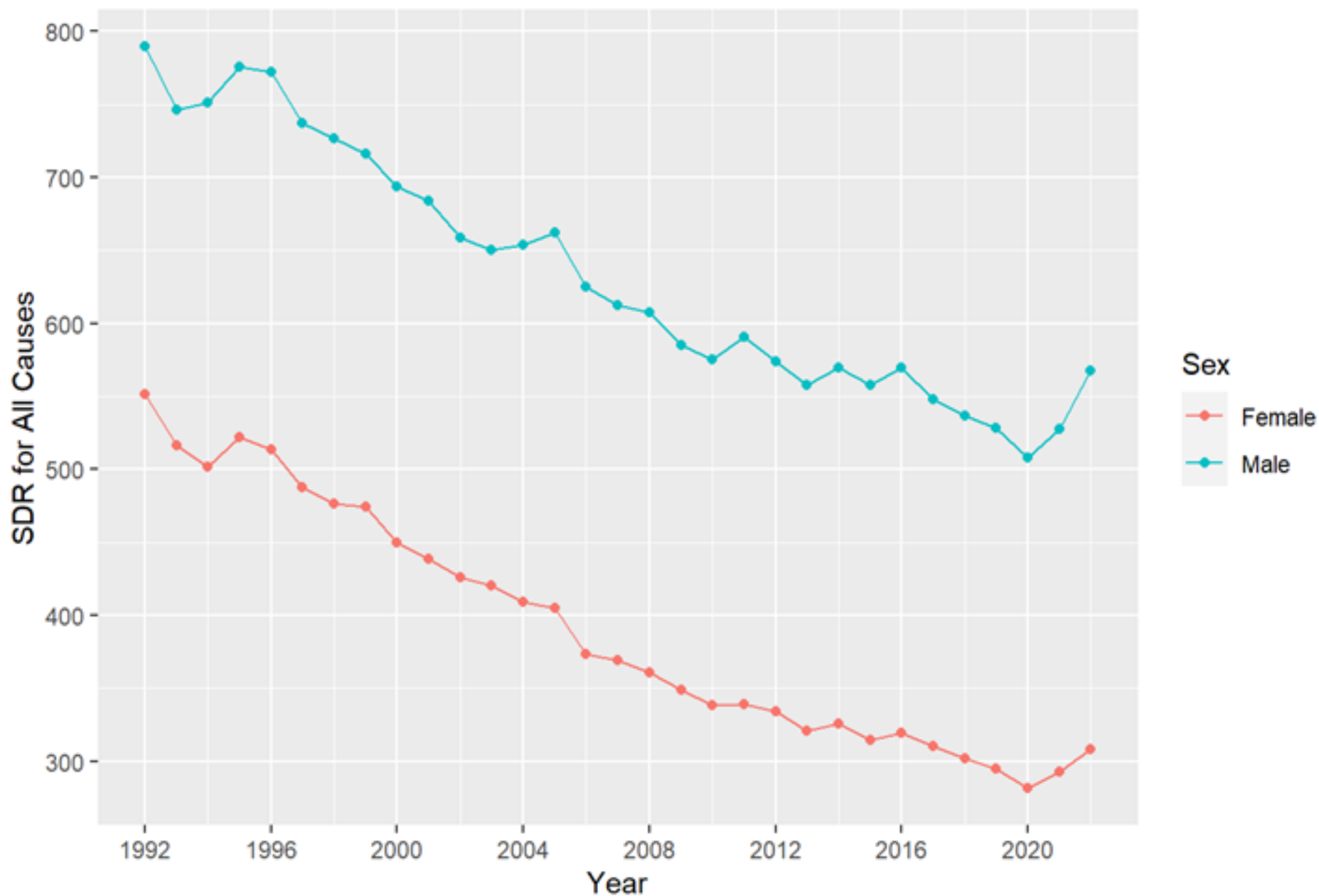
- Exploratory data analysis (EDA) is an approach to analyzing data sets to summarize their main characteristics ... EDA is for seeing what the data can tell us beyond the formal modeling. ---Wikipedia



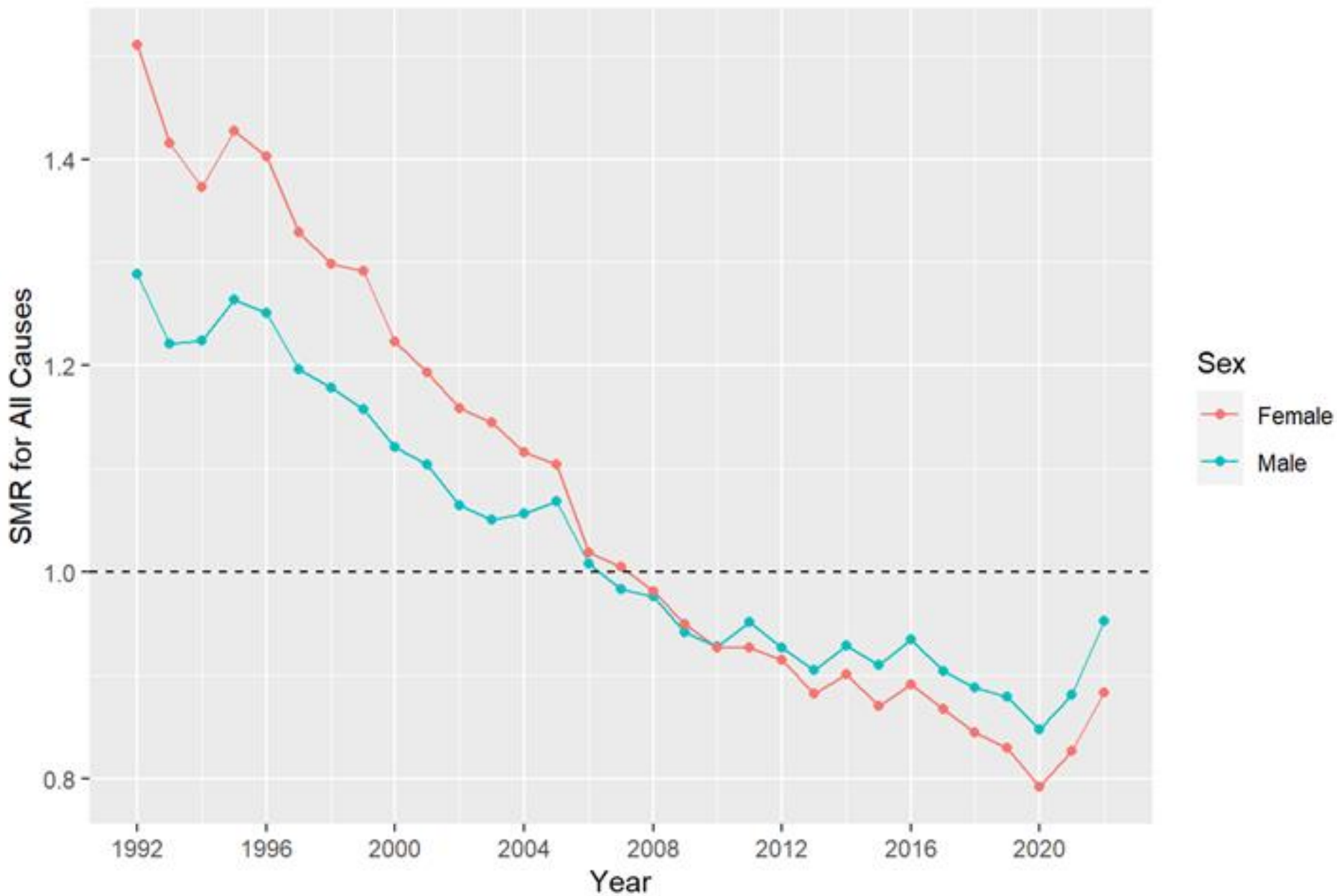
https://www.google.com/url?sa=i&url=https%3A%2F%2Fwww.aiche.org%2Facademy%2Fwebinars%2Fapplied-statistics-exploratory-data-analysis&psig=AOvVaw36ZuxAJqz27dLqU5IFzBMO&ust=1570108849384000&source=images&cd=vfe&ved=0CAIQjR_xqFwoTCJC1qLXV_eQCFQAAAAAdAAAAAAAJ

BAJ

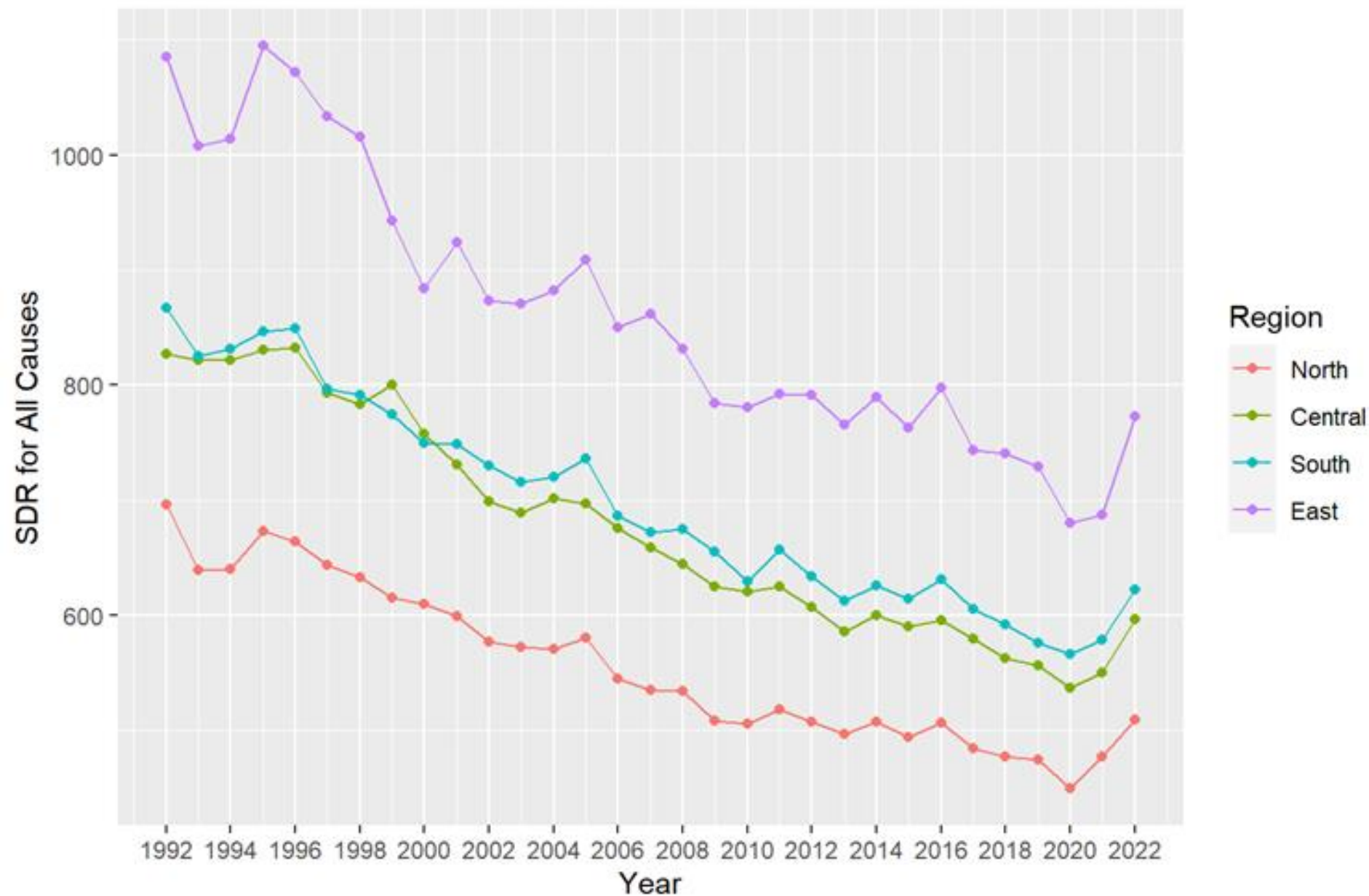
SDR Trend of Taiwan's Men and Women



SMR Trend of Taiwan's Men and Women



Regional SDR Trends of Taiwan's Male



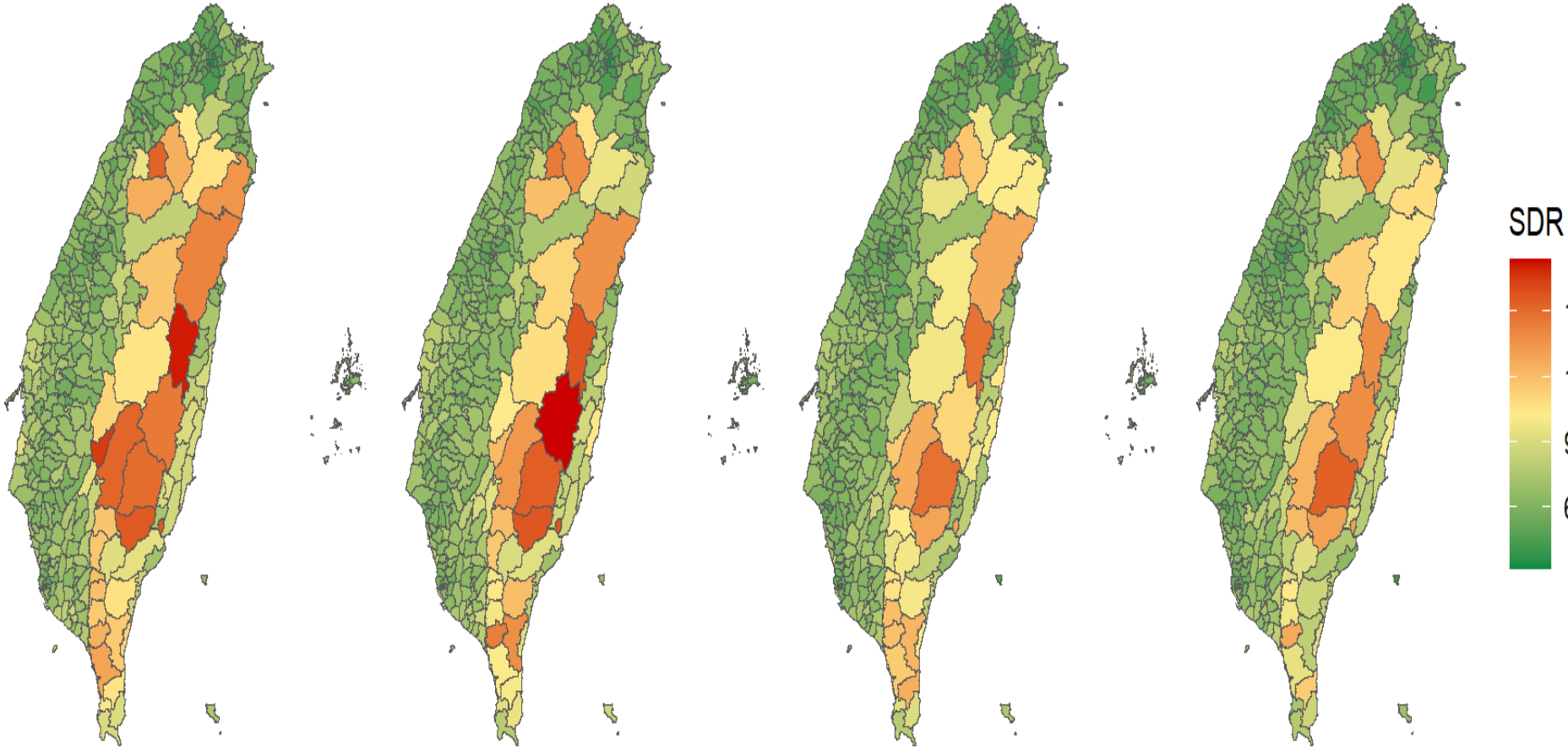
Higher SDR in Taiwan Mountain Areas

2011~2013

2014~2016

2017~2019

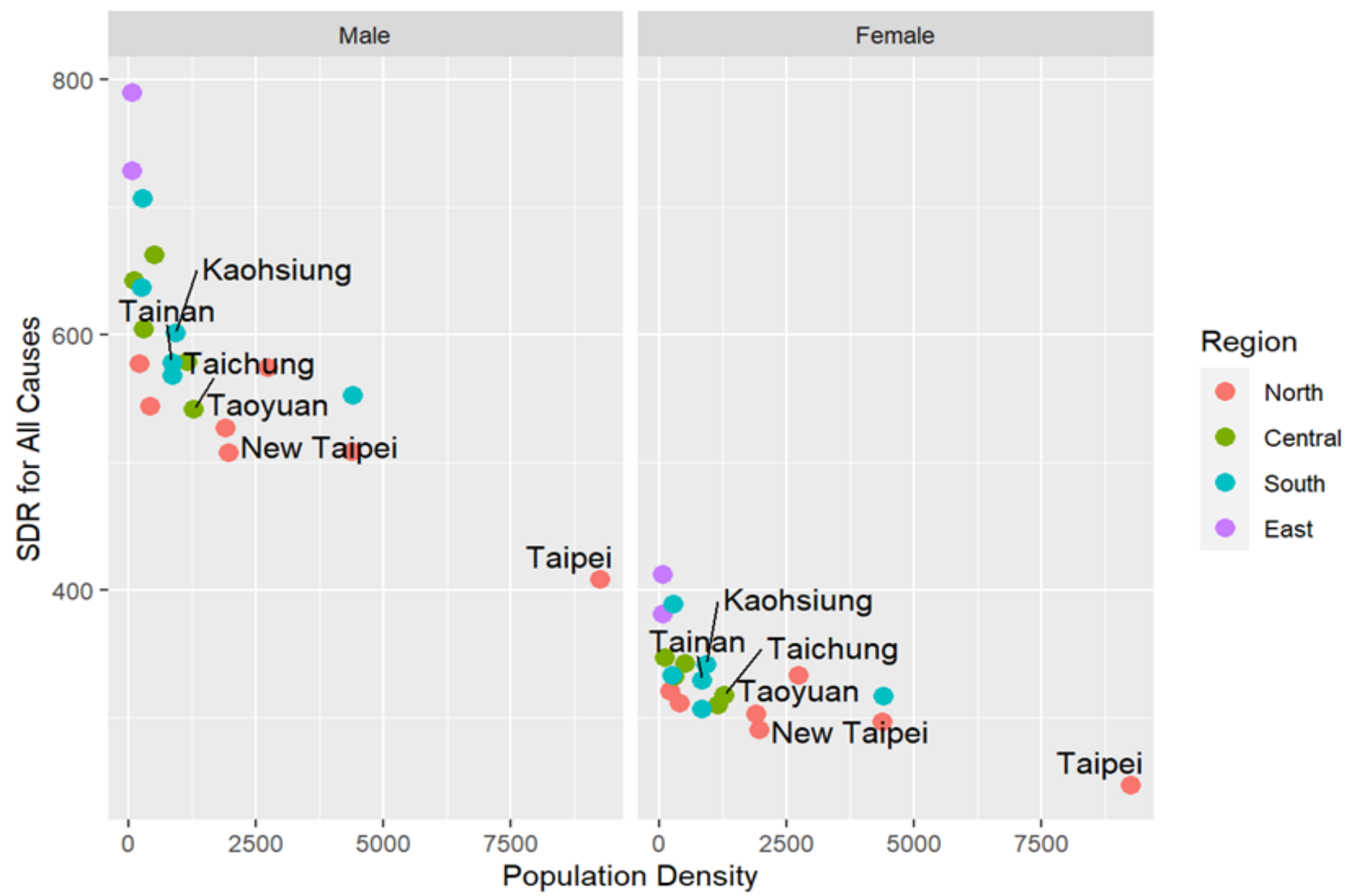
2020~2022



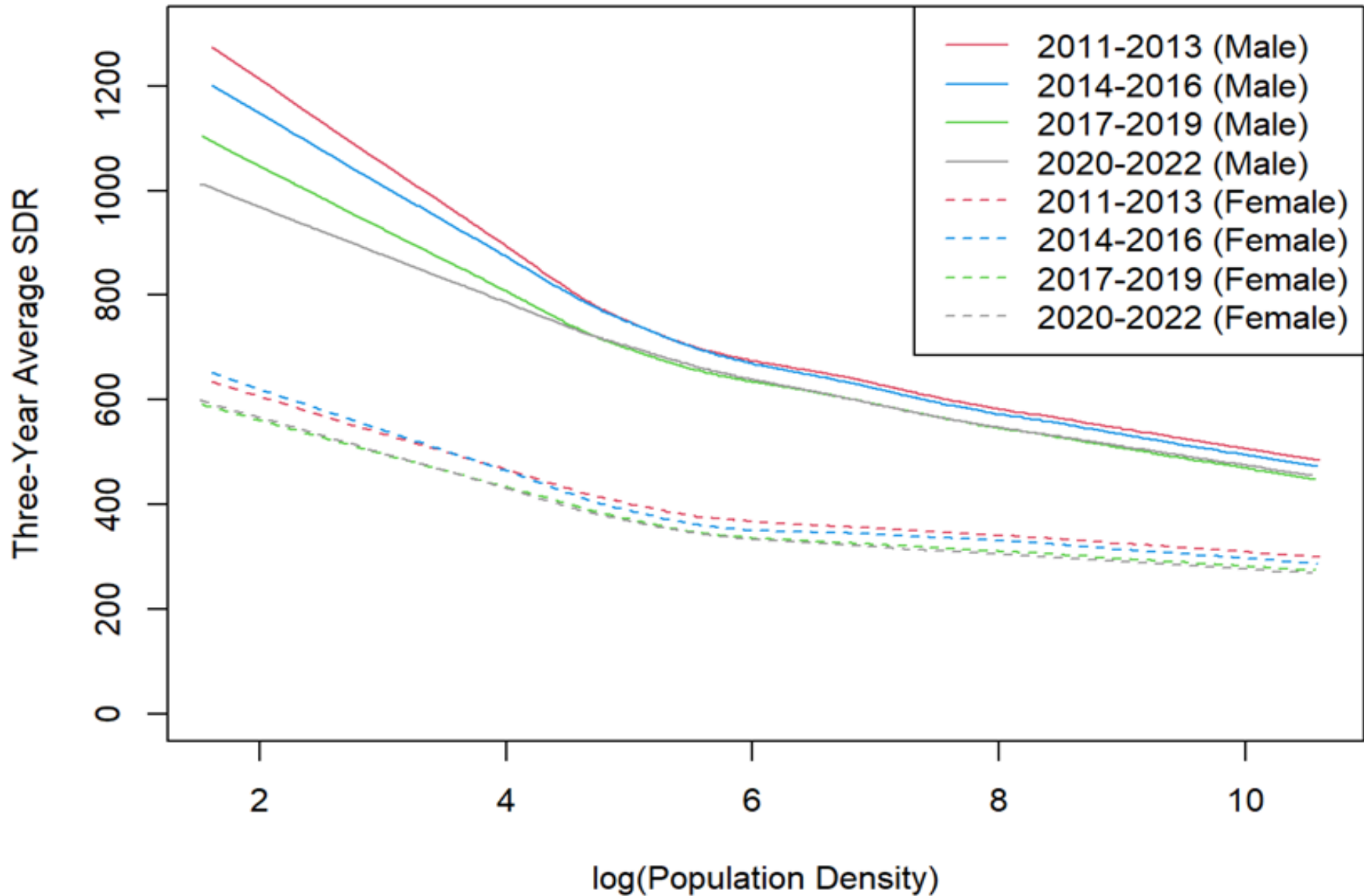
Analysis of All-cause Mortality Rates

- It seems that mortality reduction (all-cause SDR) continues in all areas of Taiwan with similar pace.
 - Women have larger improvement.
- Regional difference reduces but still exists.
 - Higher SDR and improvement rates in (mountain) areas with lower population density.
 - It seems a single LC model is not appropriate and we need more than one LC model.

SDR vs. Population Density (2011-22 Average)



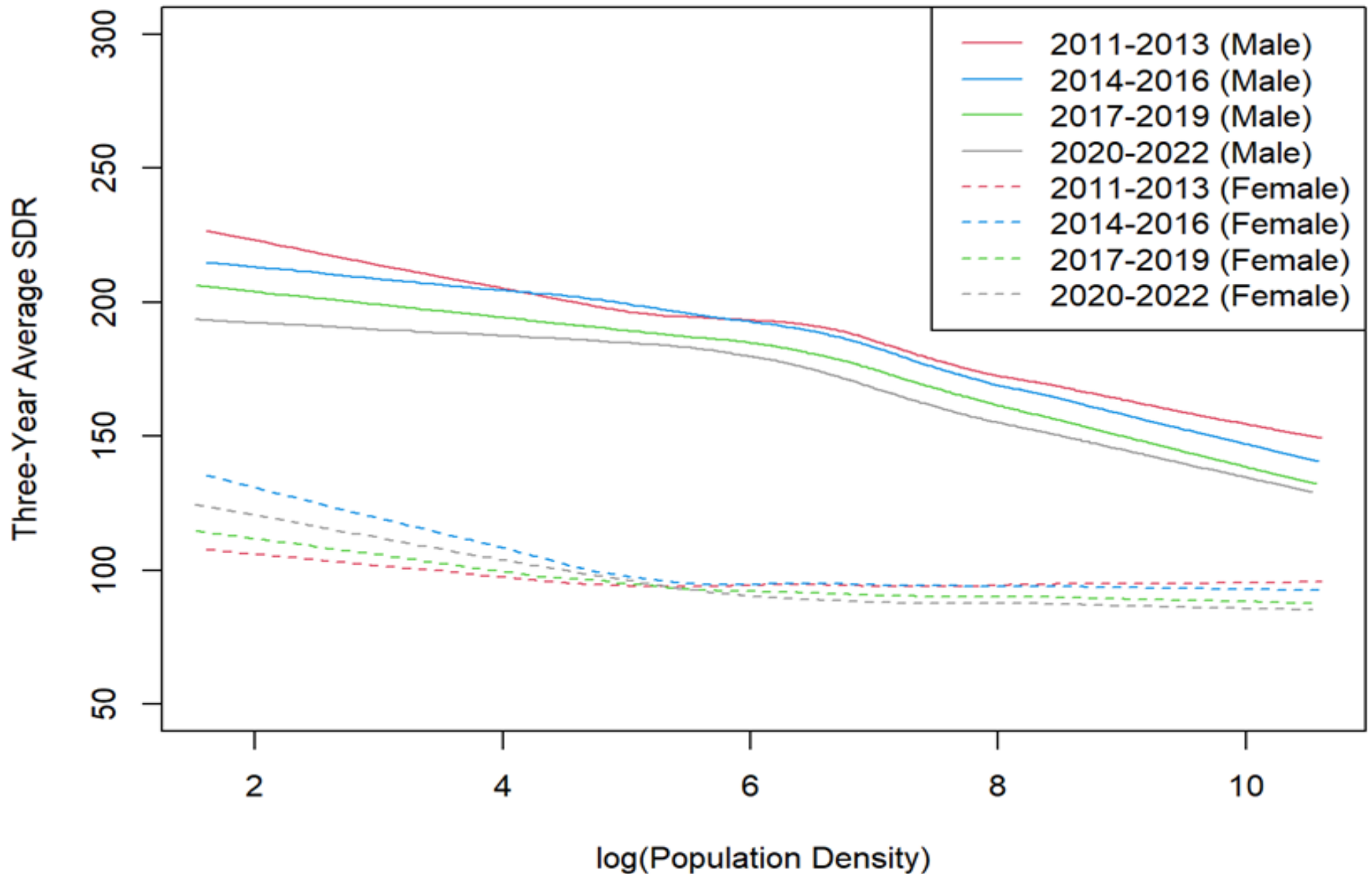
All-cause SDR vs. Population Density (2011-22)



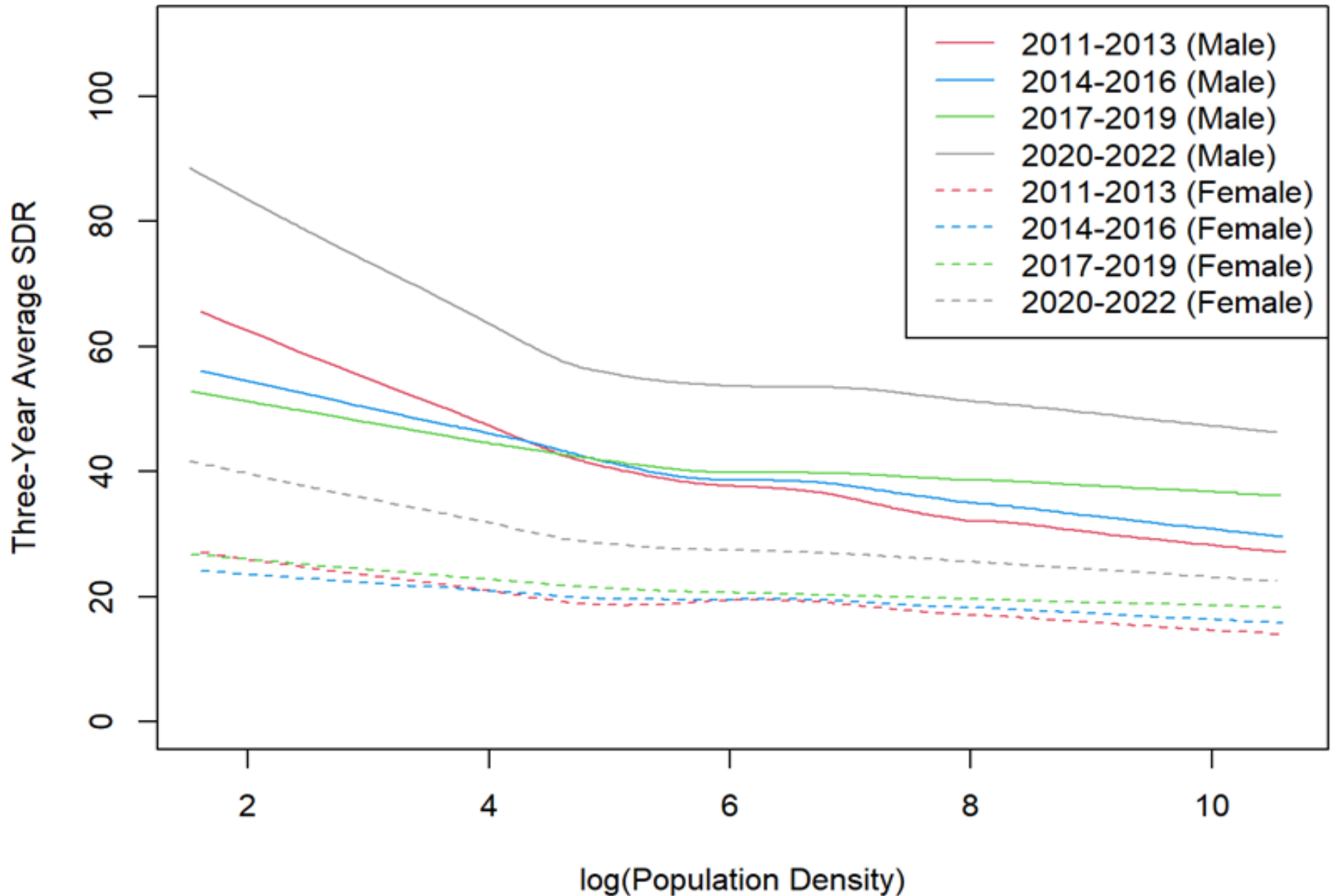
Explore Main Cause Mortality Rates

- Regional differences also appear in leading causes of death, but their trends are very different.
 - Higher SDR for areas w/ lower population density.
- Top 3 Causes: Cancer, Heart Diseases, Pneumonia
 - Cancer mortality rates decreases with time.
 - Pneumonia mortality rates increases with time.

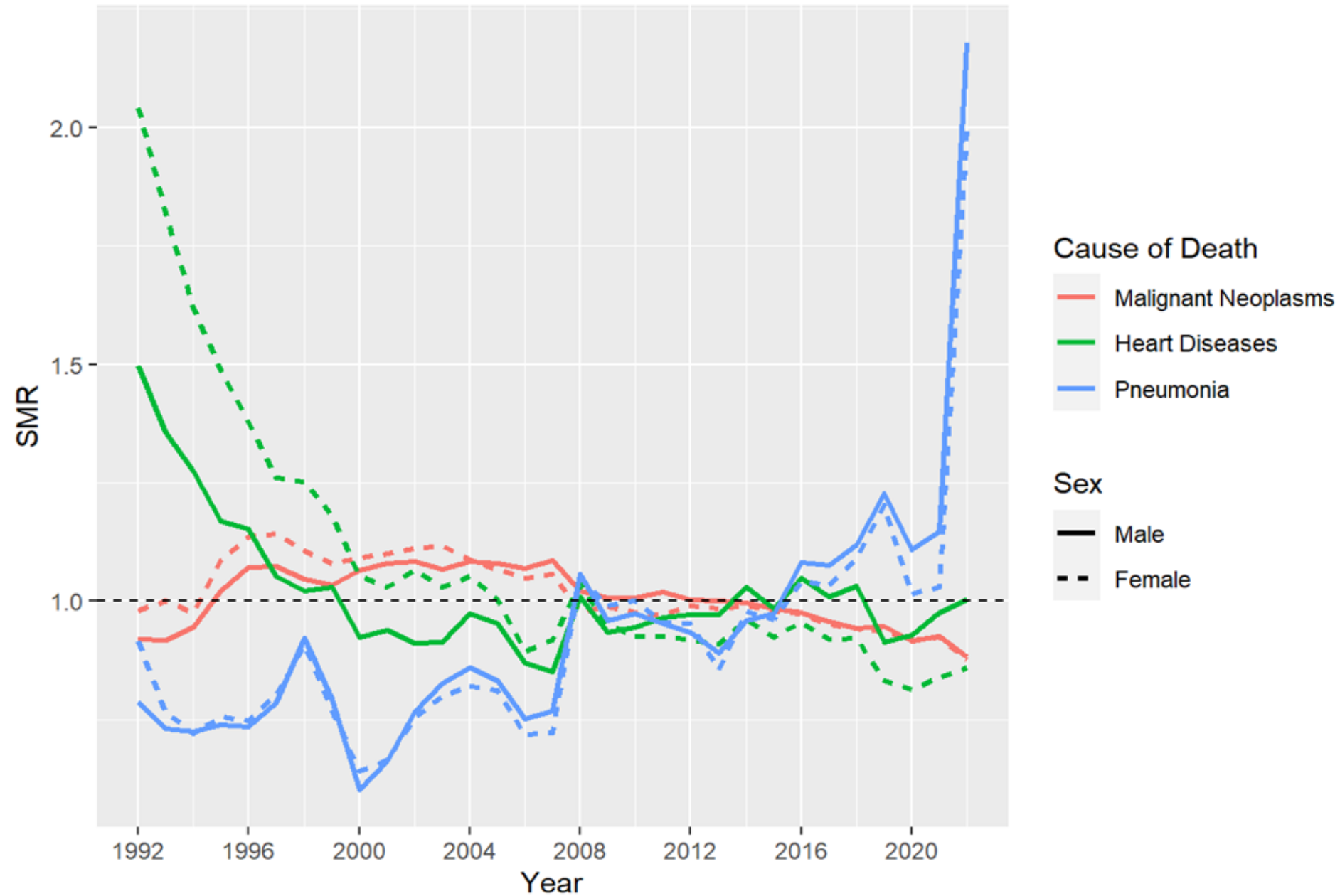
Trend of Cancer SDR (Taiwan, 2011-22)



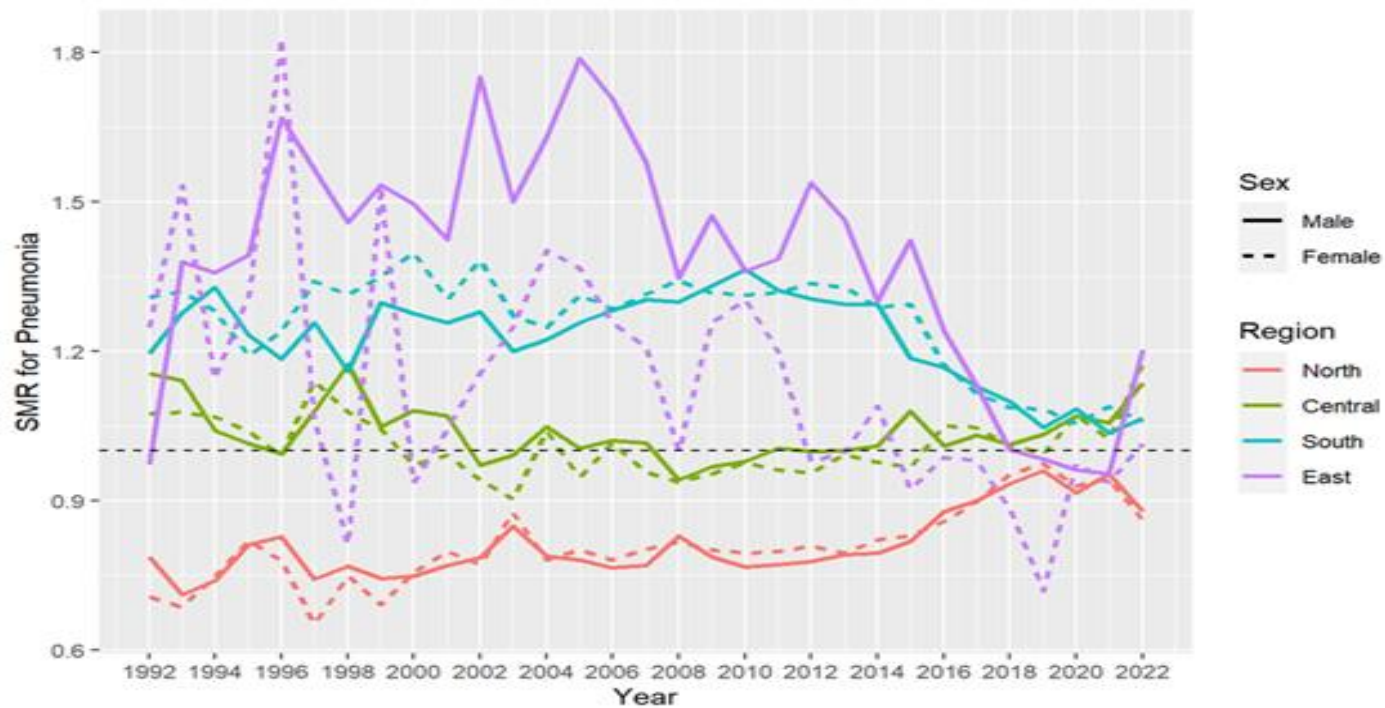
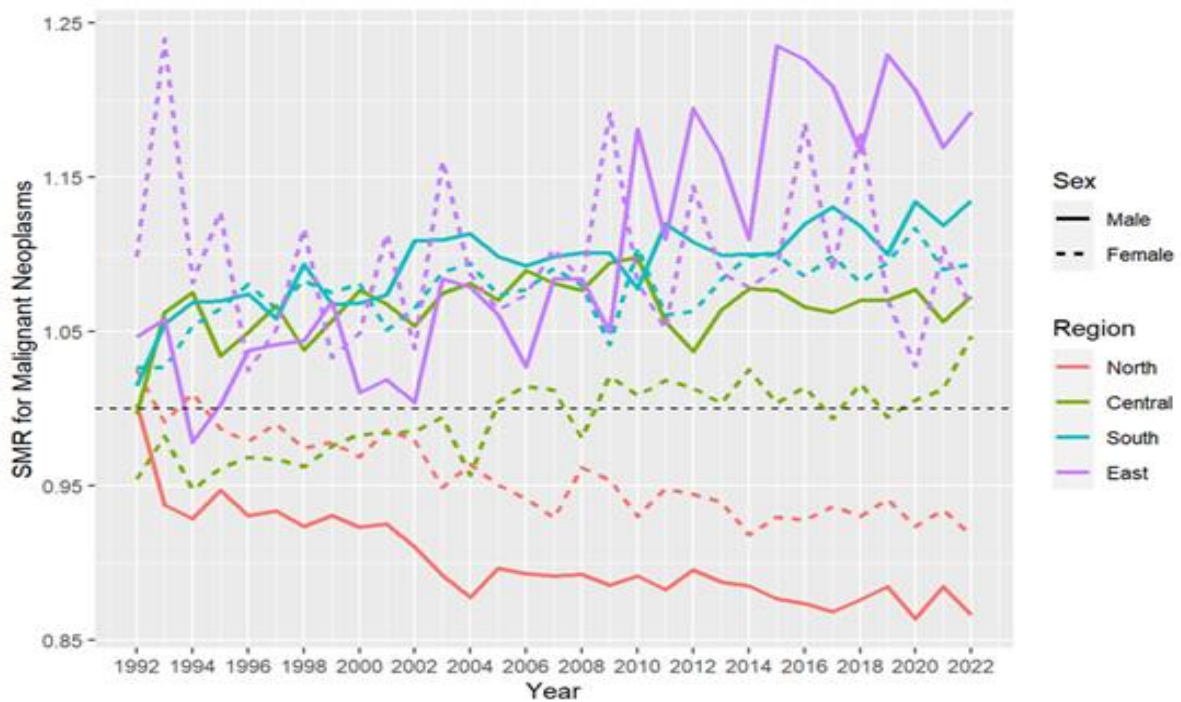
Trend of Pneumonia SDR (Taiwan, 2011-22)



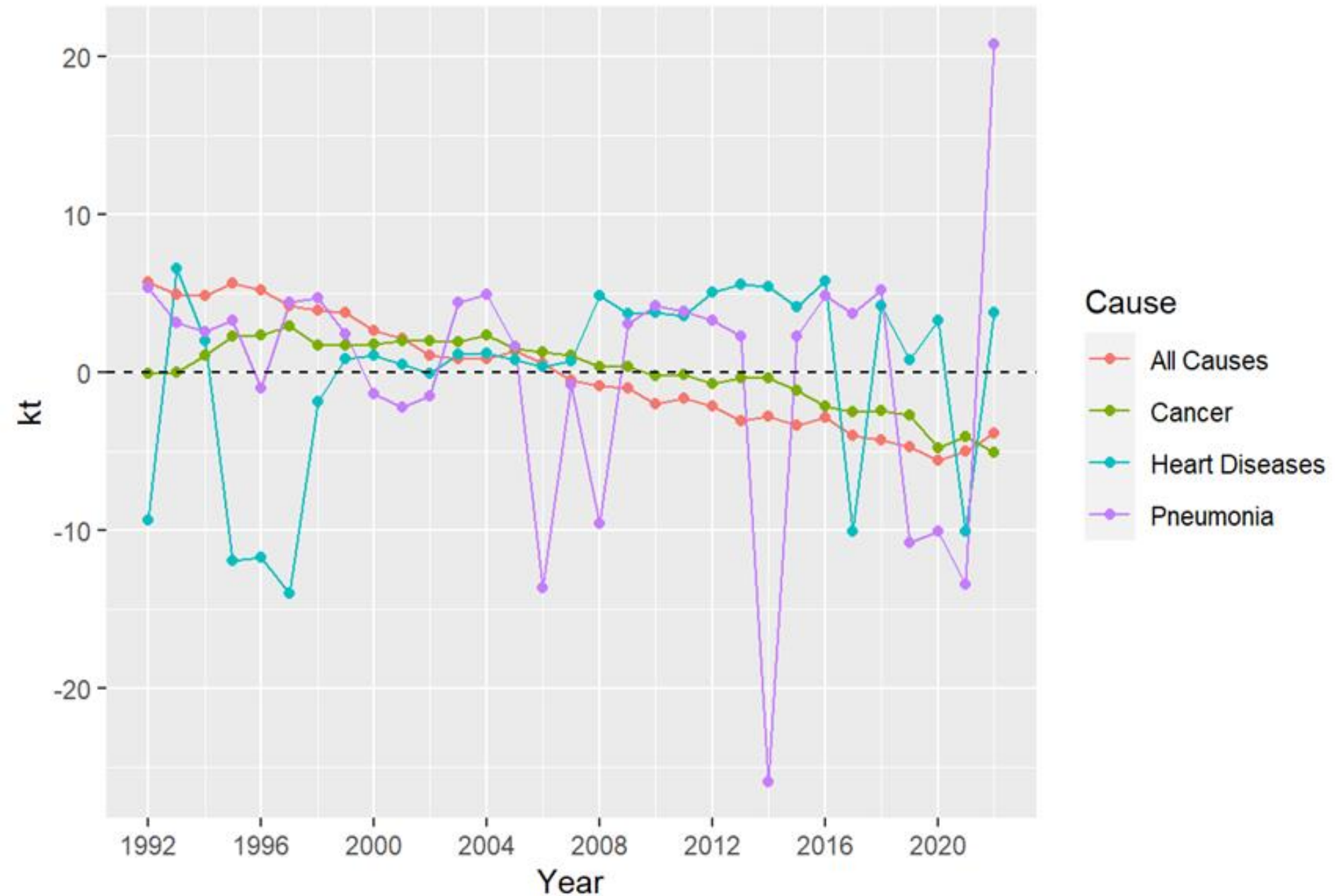
Trend of Main Death Causes SDR (1992-2022)



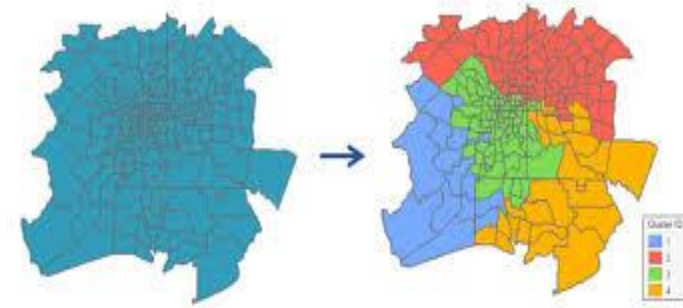
Regional Disparity of Cancer and Pneumonia SDR (1992-2022)



κ_t of LC Model for Main Death Causes



Spatial Heterogeneity



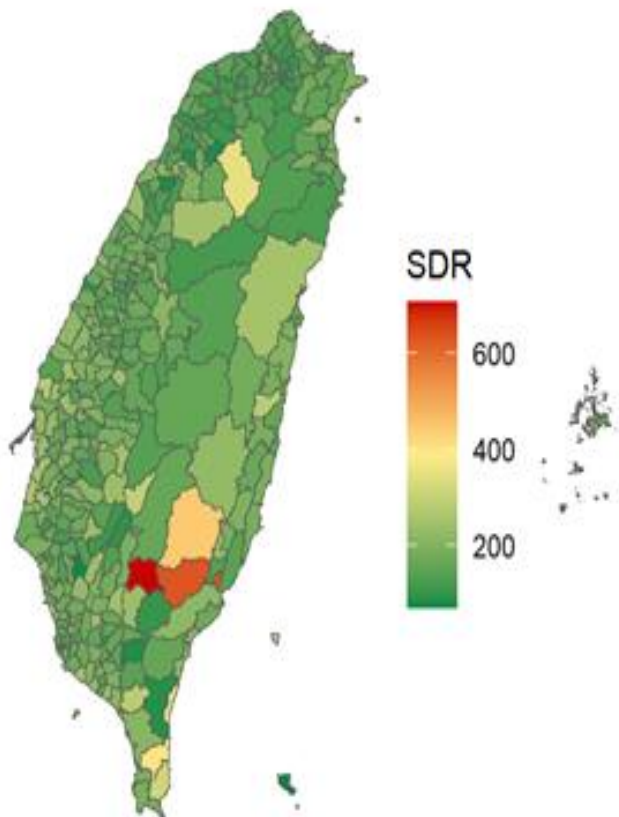
<https://pro.arcgis.com/en/pro-app/latest/tool-reference/spatial-statistics/GUID-2BD9500A-5335-45E7-9E34-5EB2A5BDE383-web.png>

	Male		Female	
	Moran's I	P-value	Moran's I	P-value
All-cause	0.5899	<.001***	0.4032	<.001***
Cancer	0.3095	<.001***	0.1207	<.001***
Heart Disease	0.4394	<.001***	0.1337	<.001***
Pneumonia	0.3795	<.001***	0.1260	<.001***

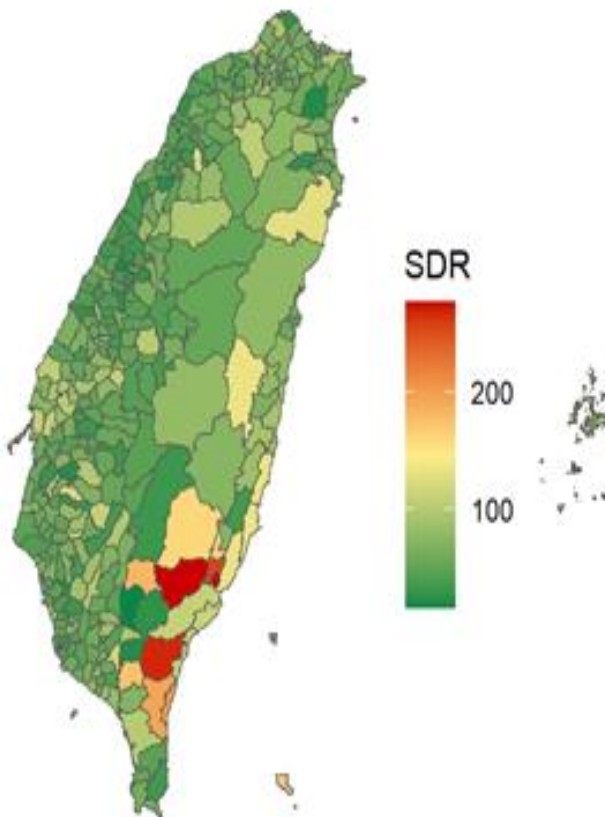
Note: Spatial clustering (cluster and/or autocorrelation)

SDR of Leading Causes (Taiwan Male, 2020)

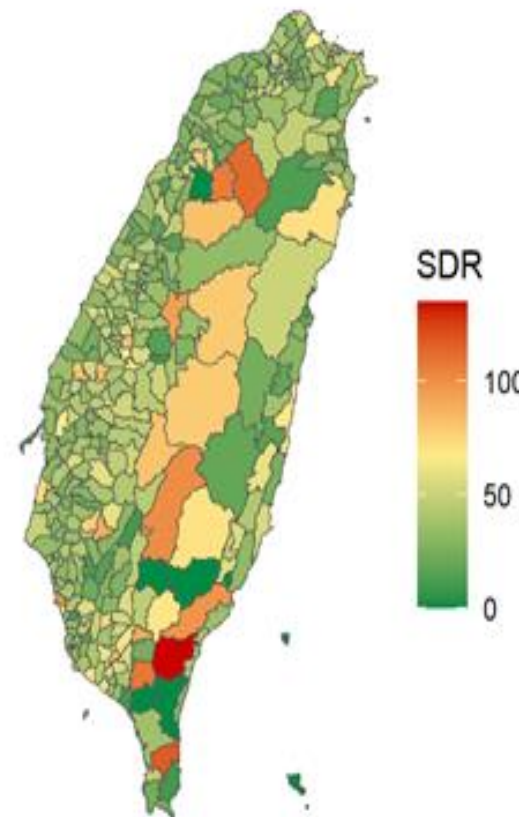
Cancer



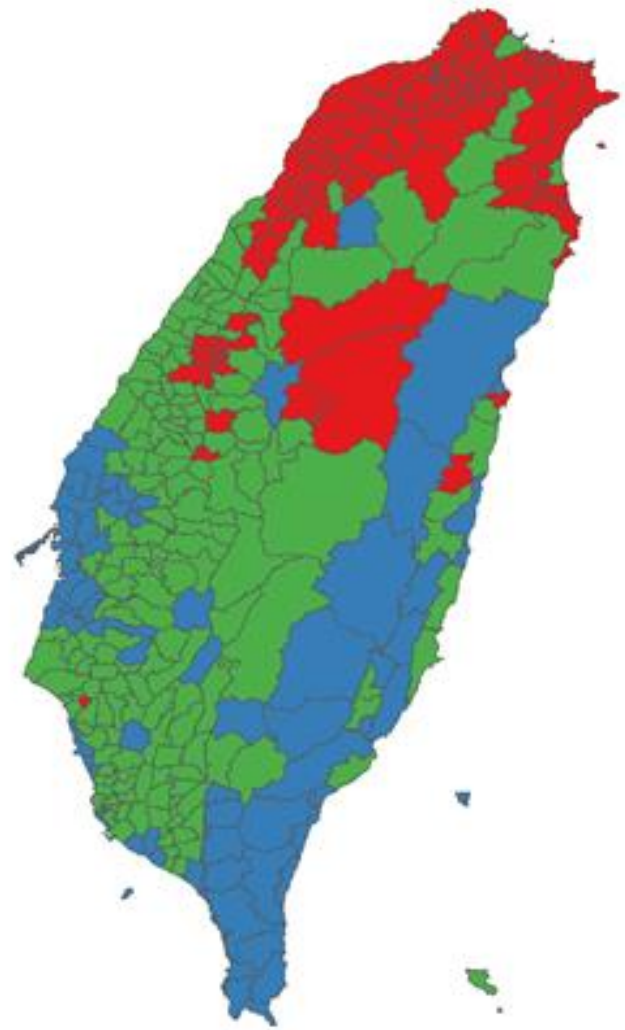
Heart Diseases



Pneumonia

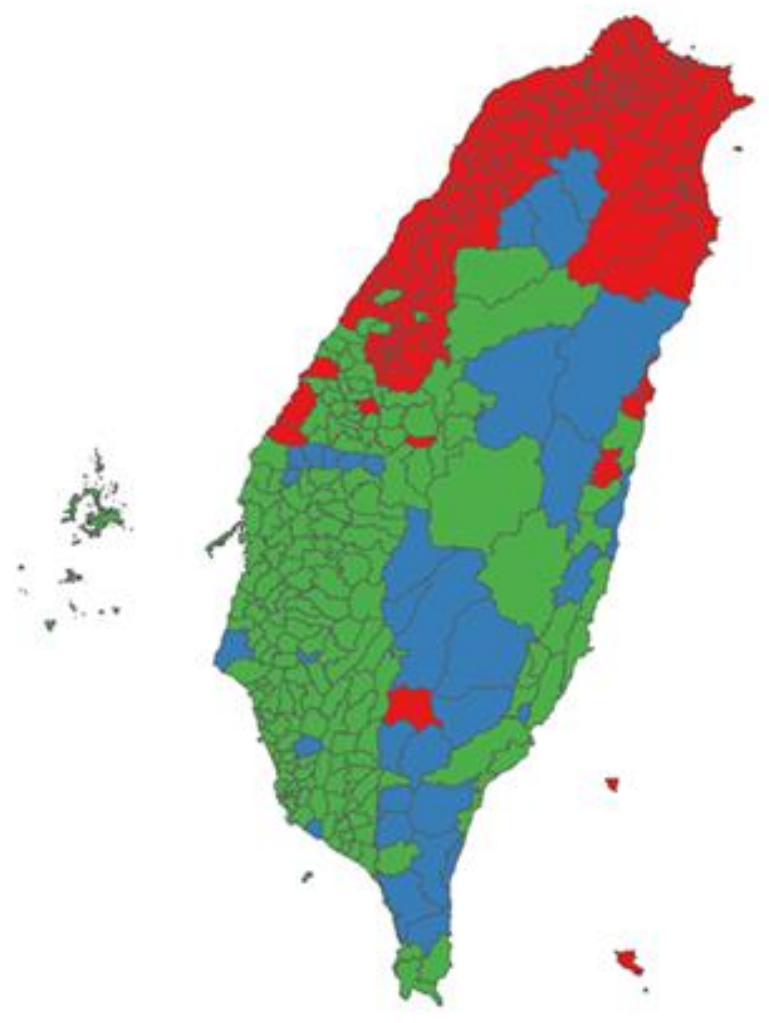


Cluster Analysis of Cancer and Pneumonia



Cluster ■ 1 ■ 2 ■ 3

Cancer



Cluster ■ 1 ■ 2 ■ 3

Pneumonia

Case 1. Ping-Tung County



- 10 years of Training data, 5 years Testing Data (1992~2022): Mortality rates of Pneumonia
→ Average testing errors of 17 combinations:

(pop. 400,000)

Approach	MAPE
Ping-Tung alone	32.58
Ping-Tung + Taiwan (PSMR)	32.55
Ping-tung + Taipei (PSMR)	32.85
Ping-tung + Cluster (PSMR)	32.18

Case 2. Tai-Tung County



- 10 years of Training data, 5 years Testing Data (1992~2022): Mortality rates of Pneumonia
→ Average testing errors of 17 combinations:

(pop. 100,000)

Approach	MAPE
Tai-Tung alone	36.90
Tai-Tung + Tai-tung all-cause (PSMR)	37.91
Tai-tung + Cancer cluster (PSMR)	40.48
Tai-tung + Pneumonia cluster (PSMR)	33.15



Conclusion

- There are regional differences in all-cause and major death cause mortality.
 - Mortality trend of different major death causes are not the same.
- The proposed small-area modification has better improvements for smaller population sizes.
 - Better results if the reference population has a similar mortality profile.

Discussions and Limitations

- Use cluster analysis (CA) to choose appropriate reference is a feasible solution.
 - Apply EDA to select variables of CA.
- Include variables other than mortality rates of major death cause.
 - e.g., Income, prop. of low-income family.
- Spatial-temporal models?

Thank you
for your Attention!

