CISNET APC Methods

COLUMBIA
Columbia University
Irving Medical Center

CIS NET

Overview

Age-Period-Cohort Effects:

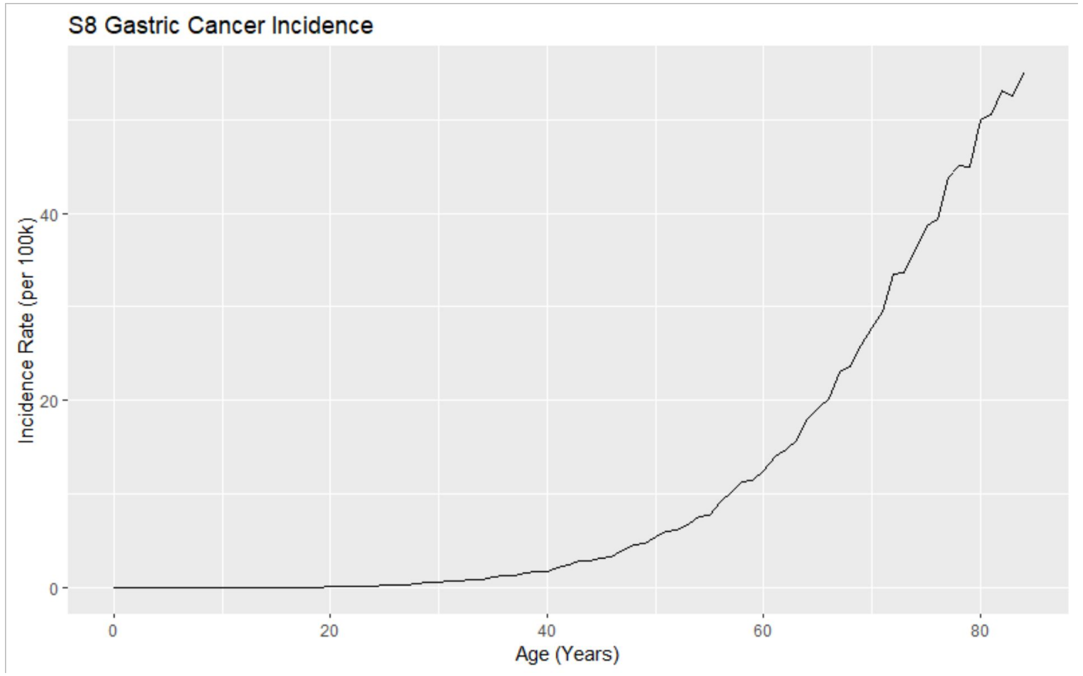
- Background
    1. Time varying effects
    2. Examples
    3. Identification Problem
    4. Existing Techniques & Modelling
- Visualization
    1. Covariate Plots
    2. Hexamaps
- Quantification
    1. Purpose
    2. Feature Selection in ML
    3. Mutual Information
    4. MI of Collinear Variables
    5. Simulated Data

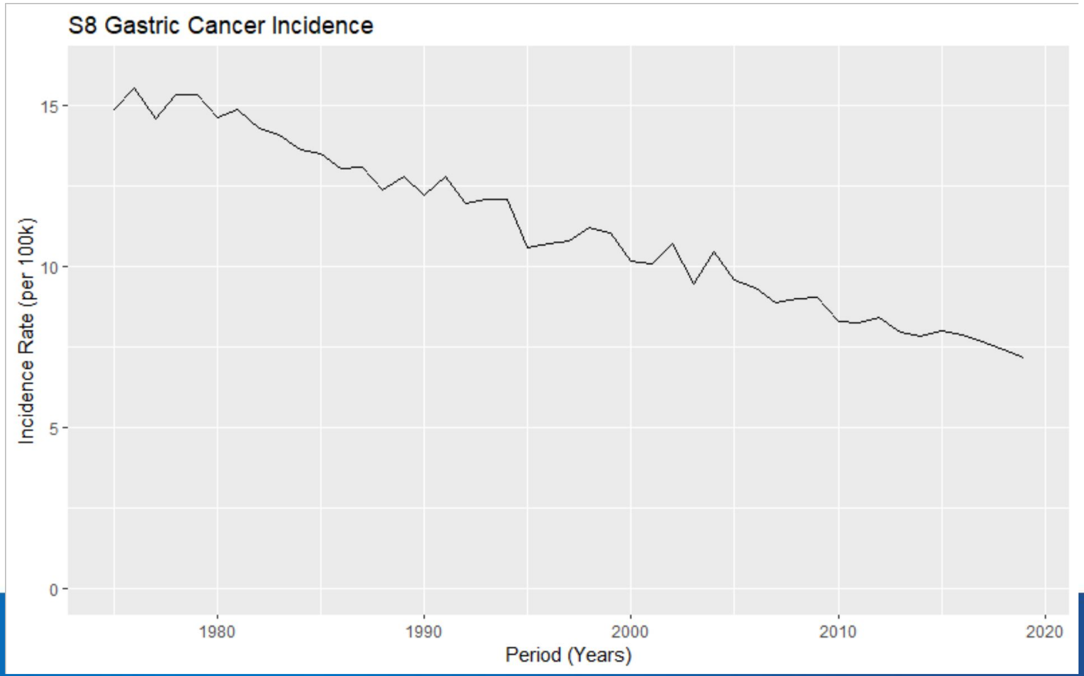
Background

i. Time varying effects
ii. Examples
iii. Identification Problem (A + C = P)
iv. Existing Techniques & Modelling

Visualization

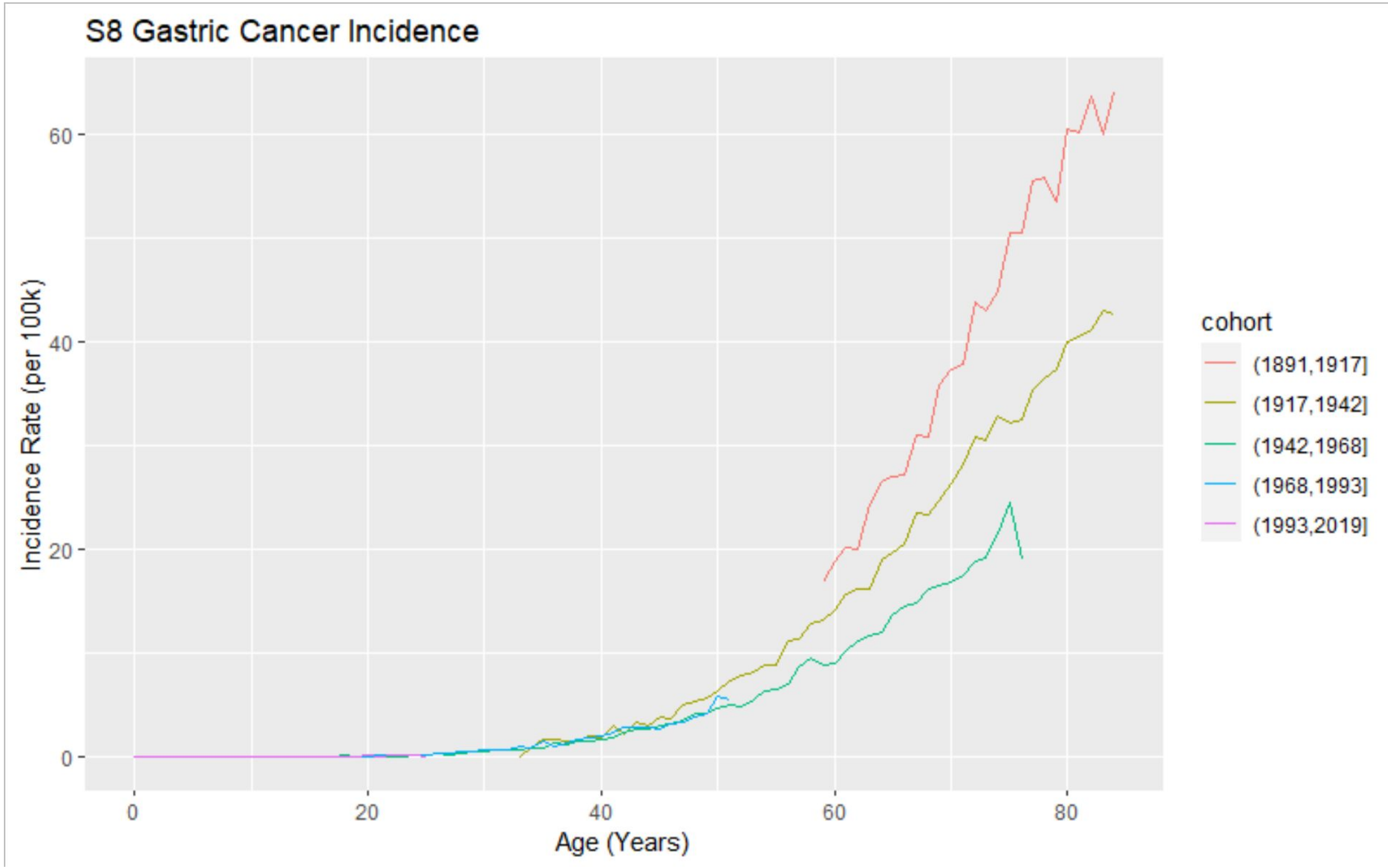
i. Covariate Plots
ii. Hexamaps





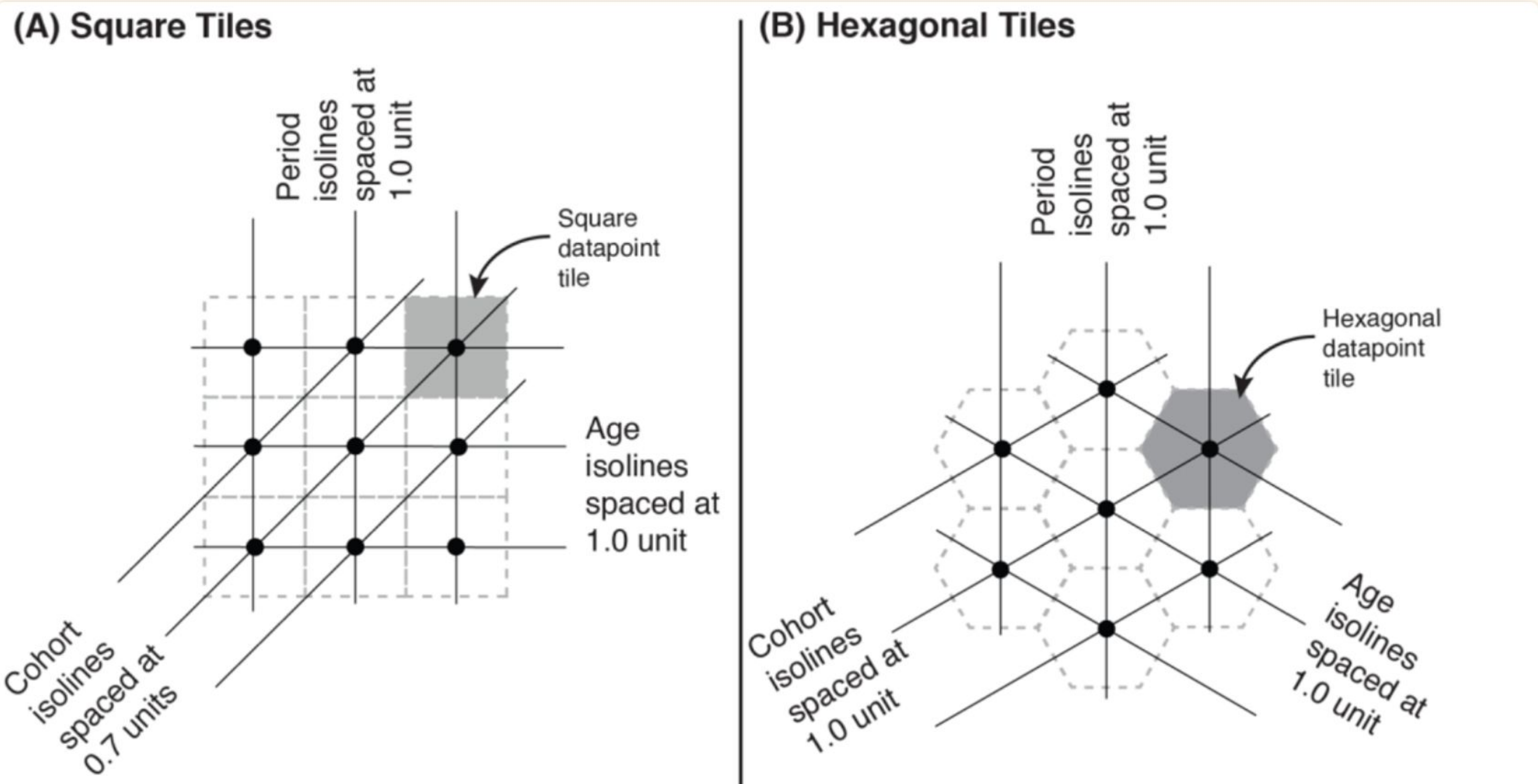
Visualization

i. Covariate Plots
ii. Hexamaps



Visualization

i. Covariate Plots
ii. Hexamaps

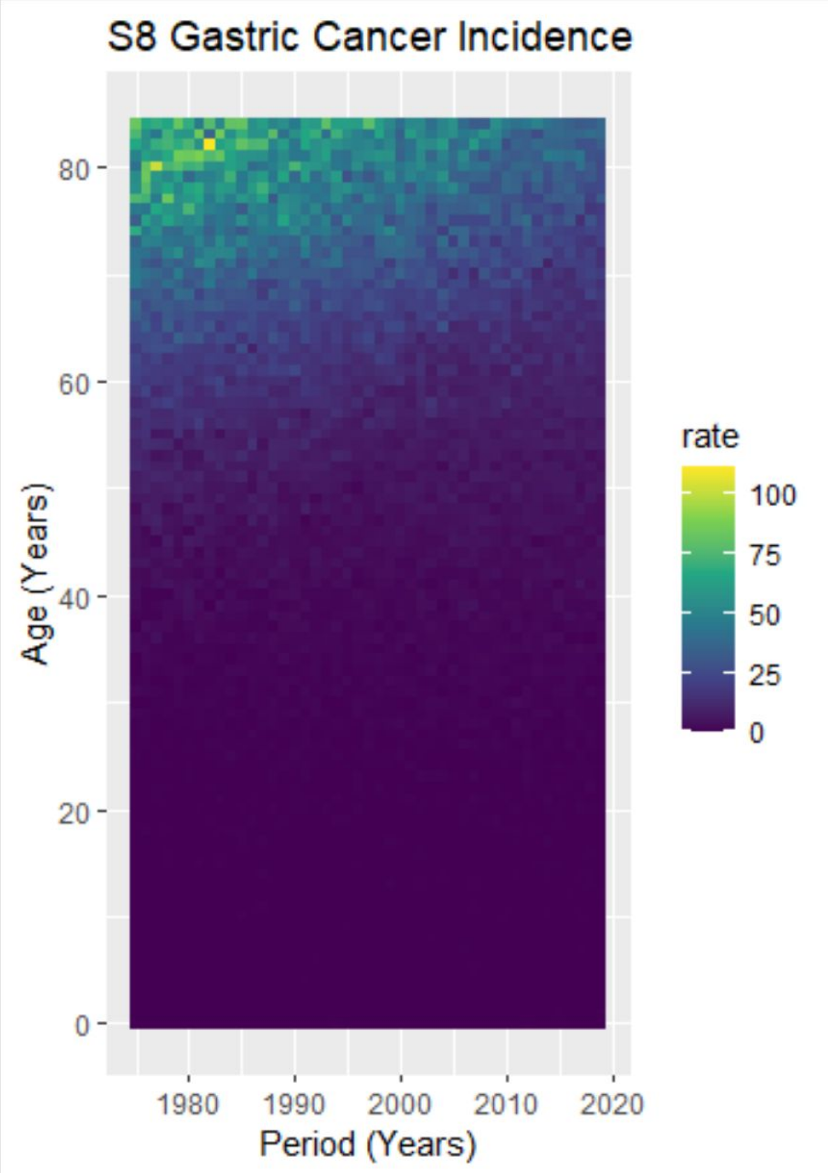


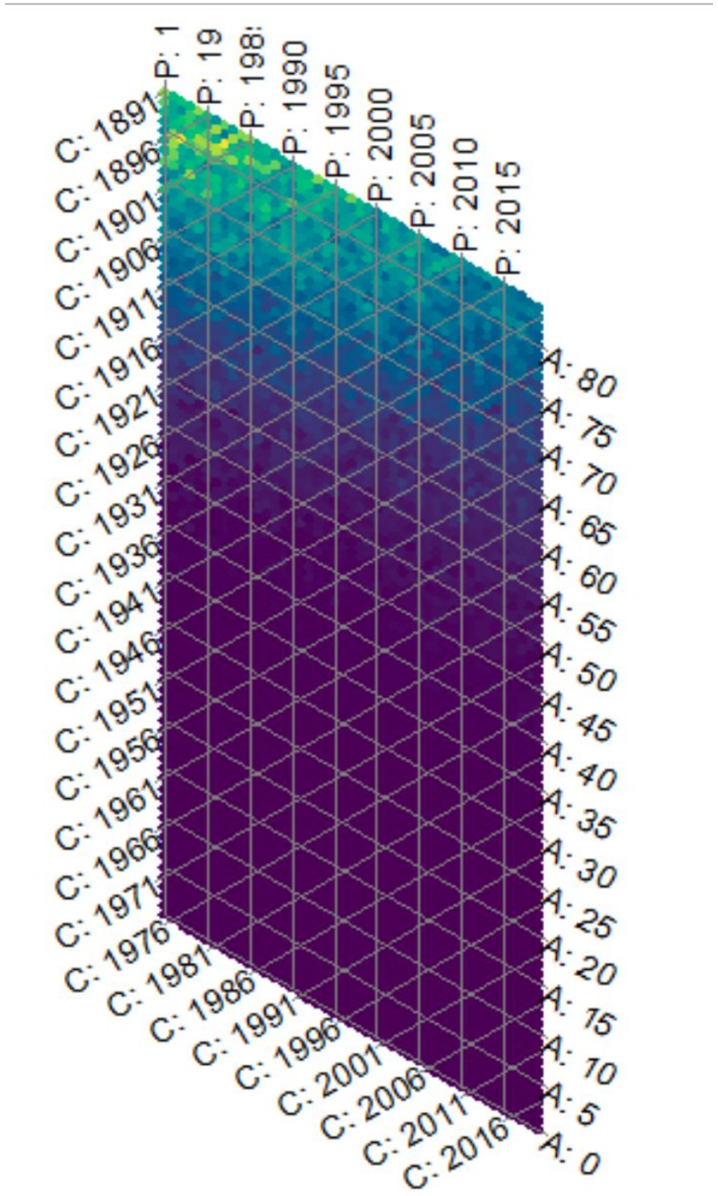
Visual distortion (right-angle triangular mesh):
1. 30% compression of cohort isolines
2. 41% expansion of the space between data tile centers along the cohort isolines
3. A single corner is shared between square tiles along the cohort axis

No distortion (equilateral triangular mesh):
1. No compression
2. Equal spacing
3. Adjacent hexagonal tiles share a full side on every axis

Visualization

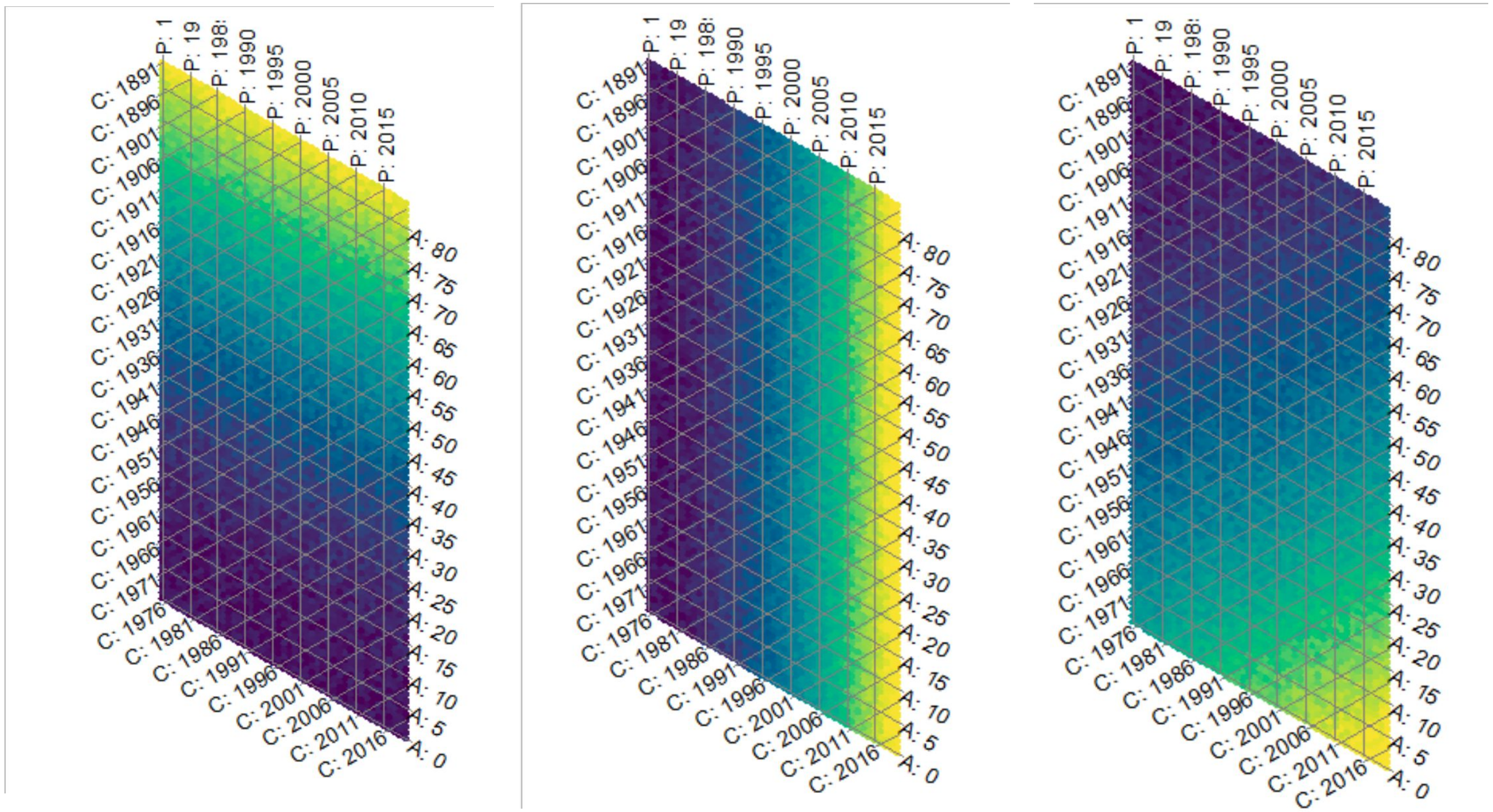
i. Covariate Plots
ii. Hexamaps





Visualization

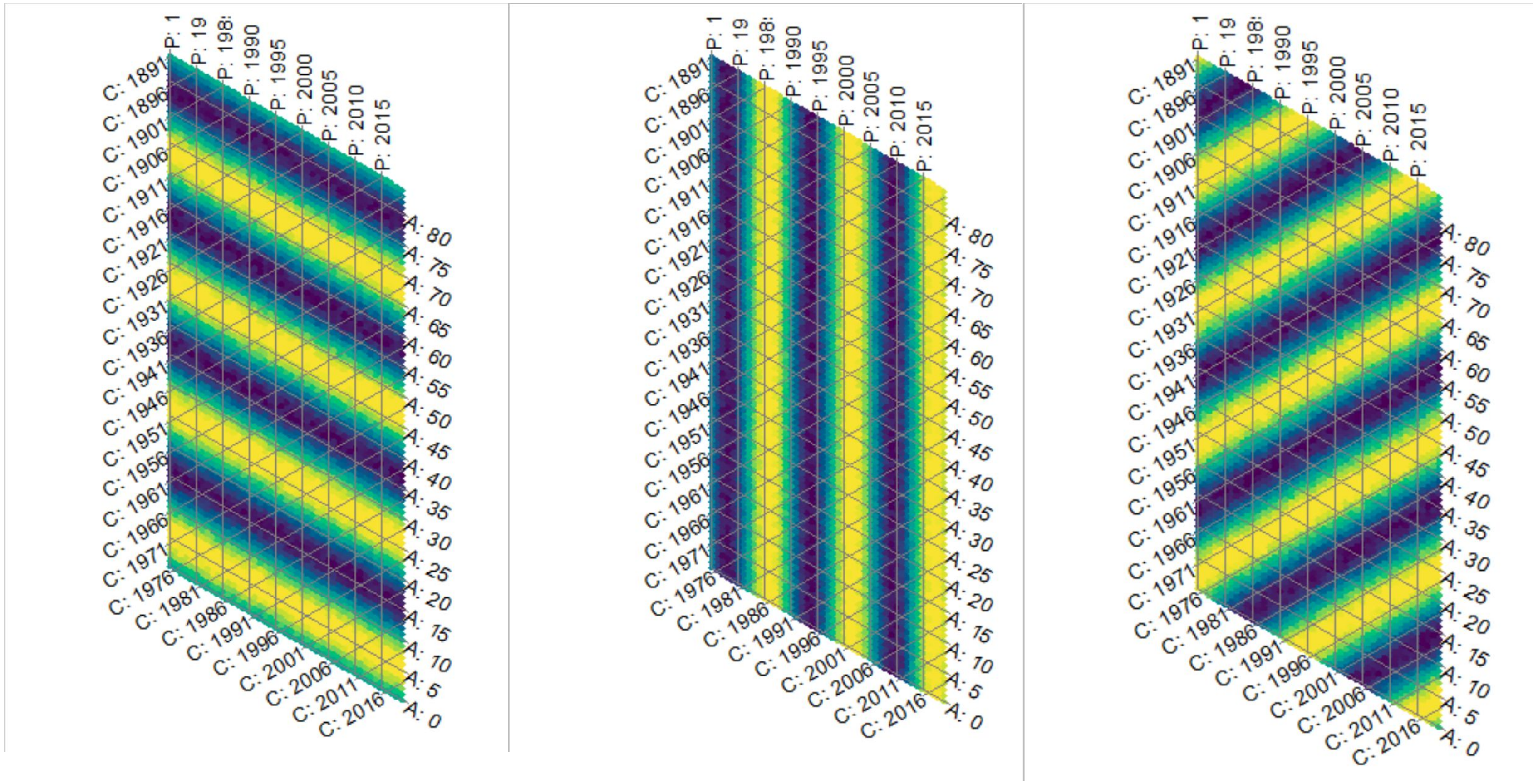
i. Covariate Plots
ii. Hexamaps



Simulated Data, Quadratic, Sinusoidal, Gaussian, Scaled to max~100, added 10% noise

Visualization

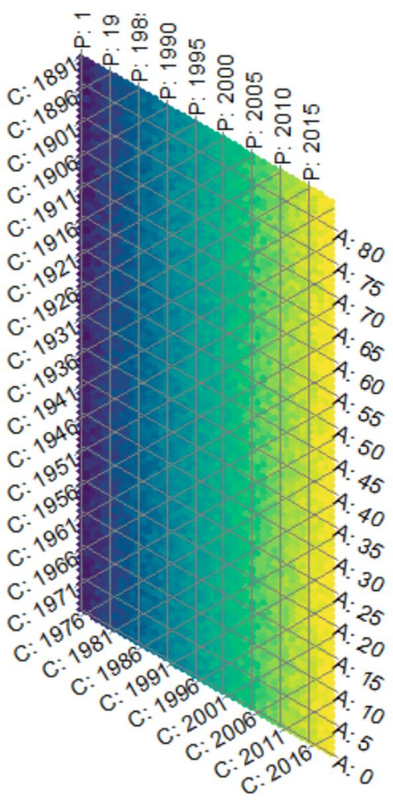
i. Covariate Plots
ii. Hexamaps

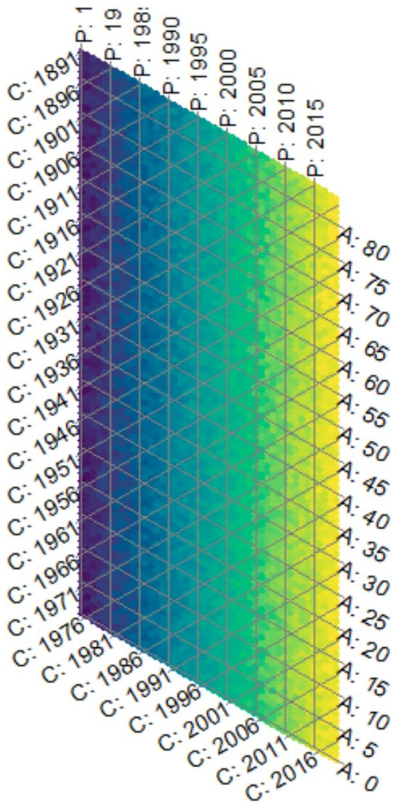


Simulated Data, Quadratic, Sinusoidal, Gaussian, Scaled to max~100, added 10% noise

Visualization

i. Covariate Plots
ii. Hexamaps

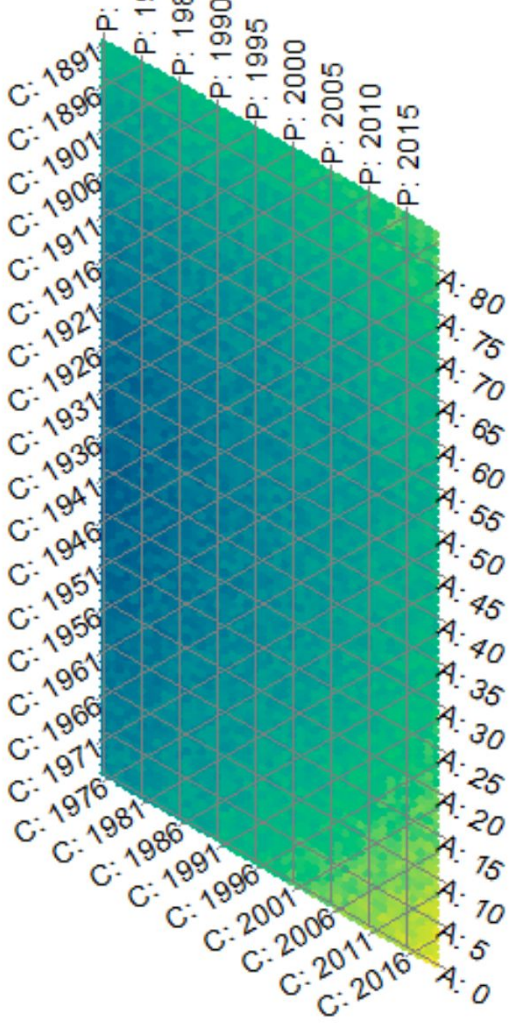


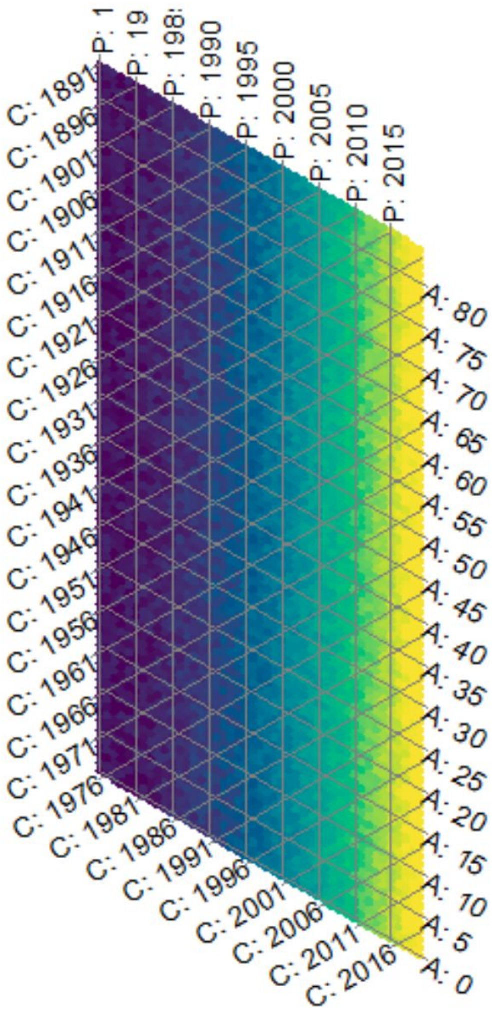


Linear Age + Cohort Mixed effect, vs Linear Period only effect Scaled to max~100, added 10% noise

Visualiza

i. Covariate Plots
ii. Hexamaps





Quadratic Age + Cohort Mixed effect, vs Quadratic Period only effect Scaled to max~100, added 10% noise

Quantification

i. Purpose
ii. Feature Selection in ML
iii. Mutual Information
iv. MI of Collinear Variables
v. Simulated Data

Quadratic	Age	Period	Cohort
Age	65.9	6.9	26.6
Period	6	79.4	11.2
Cohort	28.1	13.6	62.2

Sinusoidal	Age	Period	Cohort
Age	80.5	7	10.3
Period	8.9	84.8	10.8
Cohort	10.6	8.2	78.9

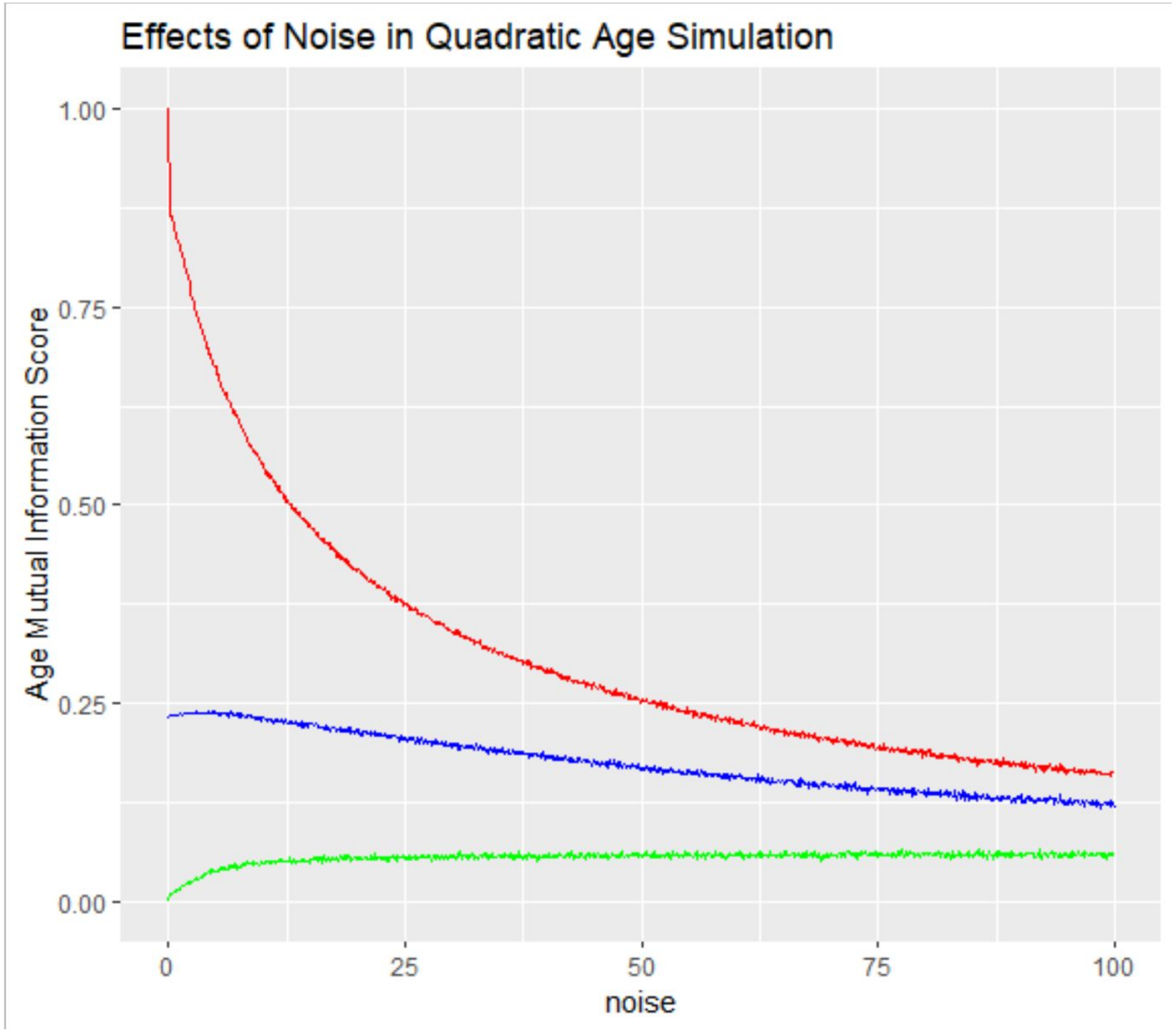
Quantification

i. Purpose
ii. Feature Selection in ML
iii. Mutual Information
iv. MI of Collinear Variables
v. Simulated Data

	Linear Age + Linear Cohort	**Linear Period**	**Quadratic Age + Quadratic Cohort**
Age	6.3	6.5	28.3
Period	79.8	79.5	34.9
Cohort	13.8	14	36.8

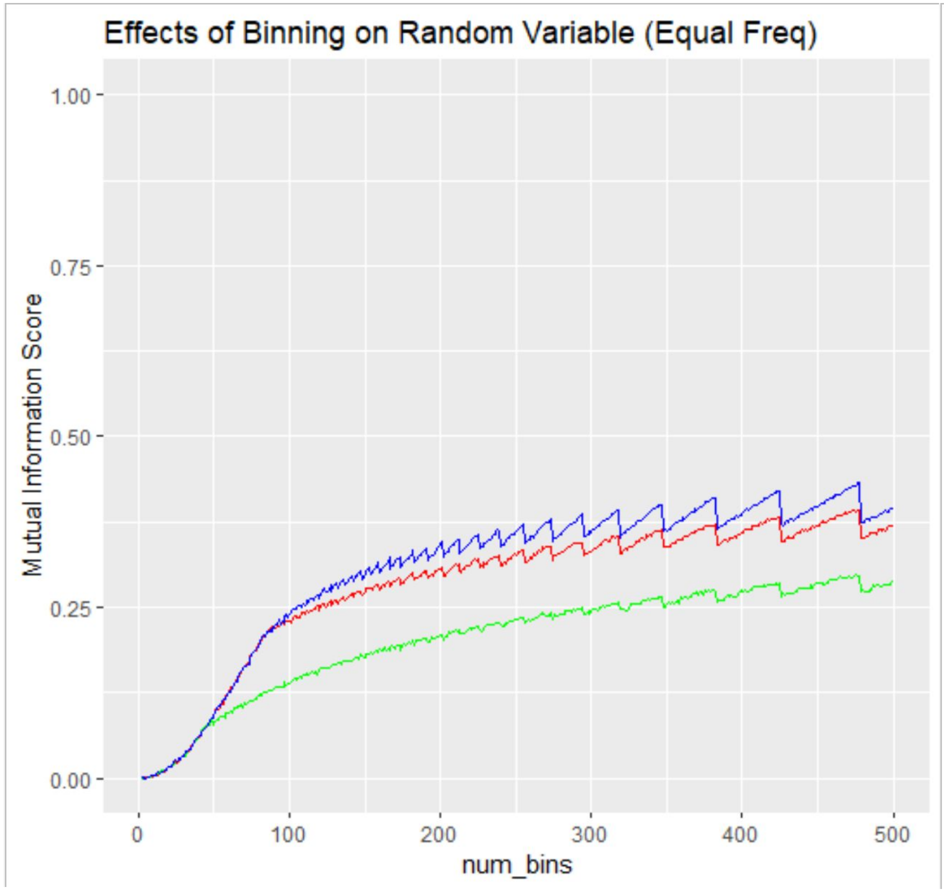
Quantification

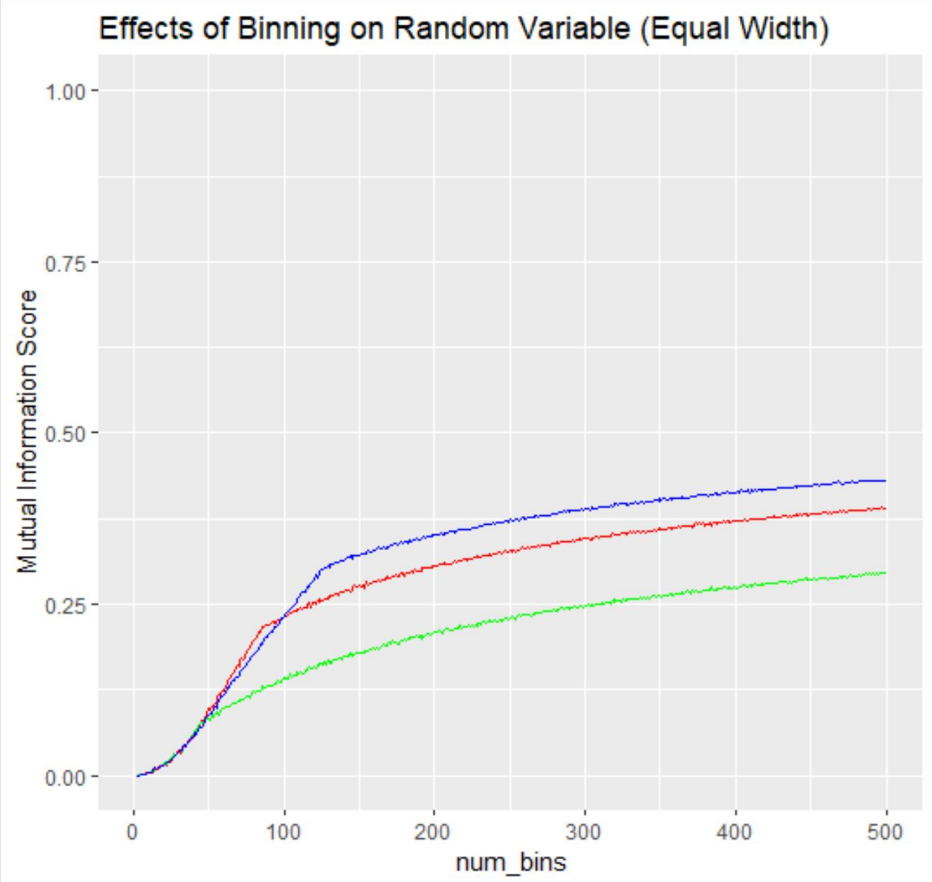
i. Purpose
ii. Feature Selection in ML
iii. Mutual Information
iv. MI of Collinear Variables
v. Simulated Data



Quantification

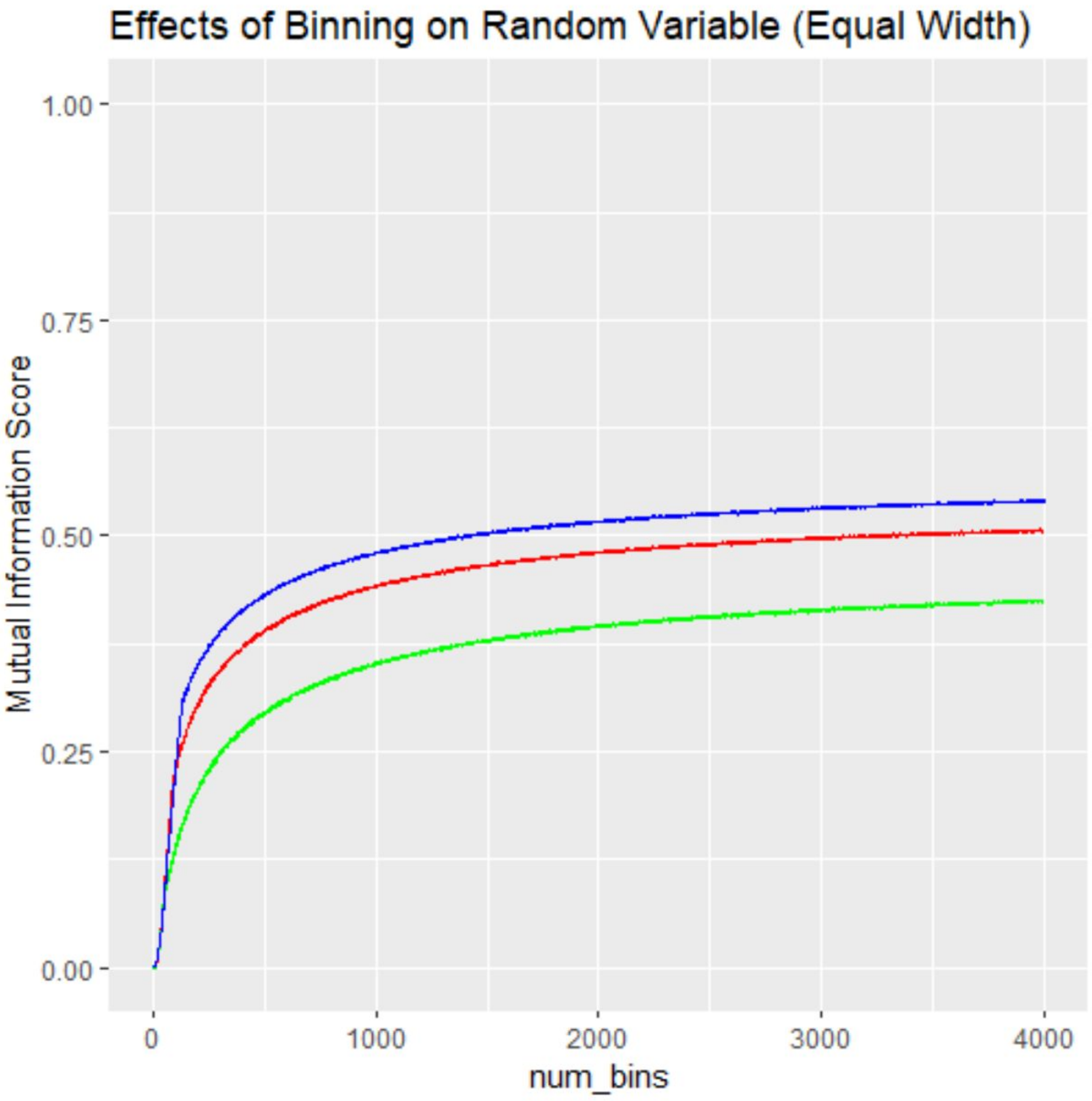
i. Purpose
ii. Feature Selection in ML
iii. Mutual Information
iv. MI of Collinear Variables
v. Simulated Data





Quantification

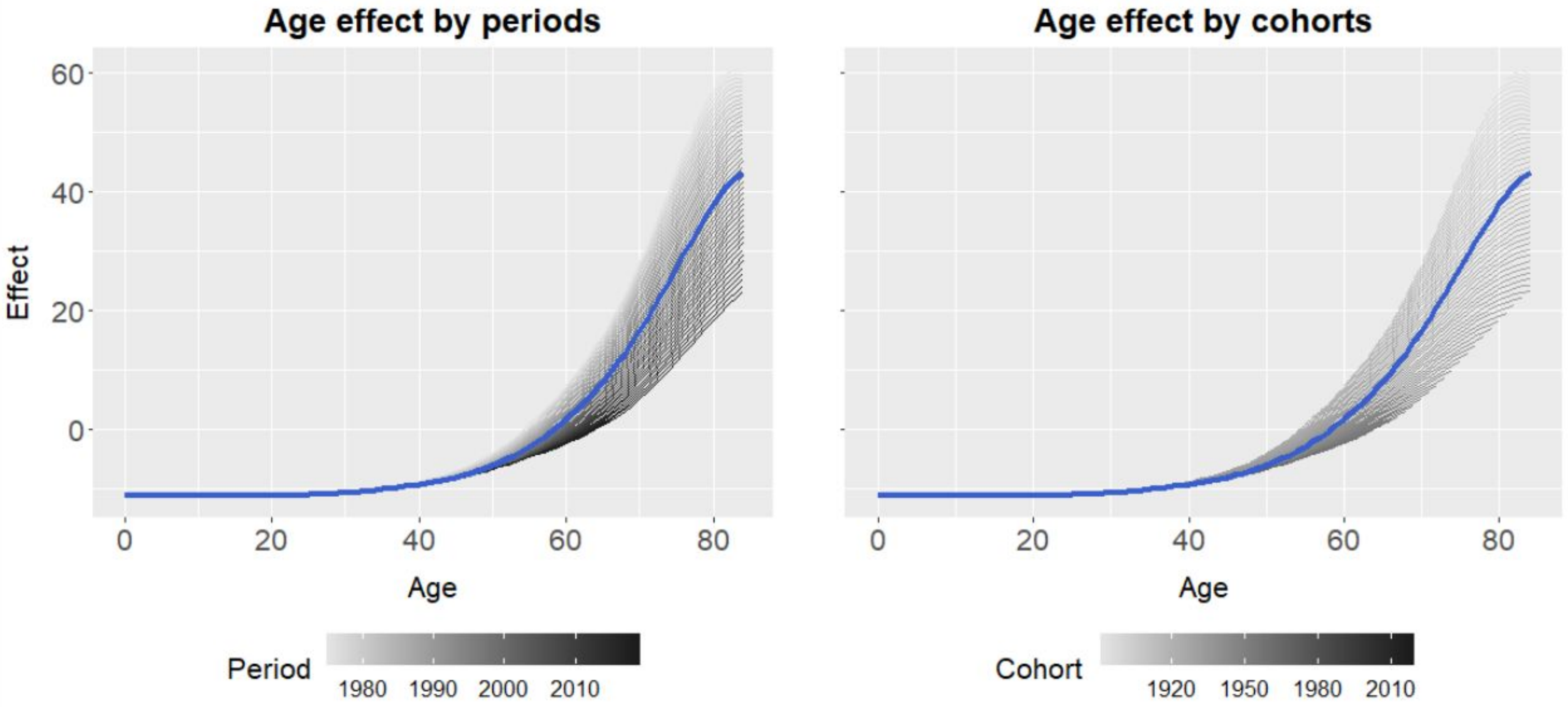
i. Purpose
ii. Feature Selection in ML
iii. Mutual Information
iv. MI of Collinear Variables
v. Simulated Data



Next Steps

Compare effect strengths with new function

- New function - model based
  i. Generative Additive Regression Models (GAMs)
    1. Assumes linear dependence on non-linear smooth functions of predictors
- Revise CI method?
  i. Currently based on 1 paper, uses population size and number of bins
  ii. Instead based on the SEER upper and lower bounds? Not straightforward
- Using Simulated Data
  i. Linear, Quadratic, Sinusoidal
  ii. A, P, C
  iii. AP, AC, PC
  iv. None, APC?
  v. Noise: 1%, 10%, 50%
- Visualization
  i. Add Partial APC Plots? (on next slide)
  ii. Hexamaps


Age effect by periods
Effect
60
40
20
0
0
20
40
60
80
Age
Period
1980
1990
2000
2010
Age effect by cohorts
0
20
40
60
80
Age
Cohort
1920
1950
1980
2010

Citations

Jalal H, Burke DS. Hexamaps for Age-Period-Cohort Data Visualization and Implementation in R. Epidemiology. 2020 Nov 1;31(6):e47-e49. doi: 10.1097/EDE.0000000000001236. PMID: 33560638; PMCID: PMC8022848.

Generalized Additive Models (GAMs)

How do they work?

- New function - model based
    i. Like a GLM but of smooth functions of predictors i.e. not limited to linear functions of predictors
    ii. Response variable can be assigned a link function e.g. log similar to logistic regression
    iii. Restrictions (based on external assumptions) on the range of smooth functions can be used to reduce overfitting (some other fancy features are in the package to reduce overfitting by penalizing ‘wiggliness’)
    iv. Recommended use of smooth tensor product transformation

Implementation

- R, mgcv package
    i. recommended and integrated with the apctools package
    ii. They cite: Weigert et al. (2021)

Examples - Weigert et al. (2021)

Semiparametric APC analysis of destination choice patterns: Using GAMs to quantify the impact of A/P/C on travel distances

- Purpose
  1. Examine how A/P/C are related to travel distances
- Data
  1. Cross sectional survey of German tourism 1971-2018 (trend towards further travel distances)
  2. Changes attributed to life cycle (age), macroeconomic developments (period) and generational membership (cohort)
  3. Uses GAMs on te(age, period) to circumvent identification problem
  4. Cohorts are represented as an interaction between age and period (treated on uneven footing)

Table 2. Overview of the travel distance categorization used in the analyses.

Travel distance	Exemplary destinations	Relative frequency
<500 km	Germany and neighboring countries	1971: 57.9% 2018: 31.9%
500–1000 km	Neighboring and close European countries	1971: 24.5% 2018: 18.7%
1000–2000 km	European countries (e.g. Spain, Portugal, Malta, and Finland)	1971: 15.3% 2018: 28.4%
2000–6000 km	North Africa, Middle East, Russia, and Mongolia	1971: 1.5% 2018: 11.5%
>6000 km	America, Africa (excluding North Africa), Asia, and Australia	1971: 0.8% 2018: 9.5%

Examples - Weigert et al. (2021)

APC Model Structure

Estimating mean marginal effects

- Visualization
    i. Using partial APC plots to display the nonlinear variation over cohorts for a specific age group
    ii. Convert effect strength to odds ratios

$$\log\left(\frac{\pi_i}{1-\pi_i}\right) = \beta_0 + f_{\text{ap}}(\text{age}_i, \text{period}_i), \quad i = 1, \ldots, n \qquad (3)$$

with π_i the probability to travel in the respective distance category, β_0 the intercept, $f(\cdot,\cdot)$ a two-dimensional nonlinear function, and n the number of individuals. In the remainders of this work, we call model (3) the *pure APC model*. All quantitative interpretations and visualizations of effects in this study are based on odds ratios ($\text{OR} = \frac{\pi_i}{1-\pi_i}$).

$$f_a(\text{age}_a) = \frac{1}{P}\sum_{p=1}^{P} f_{ap}\left(\text{period}_p \mid \text{age}_a\right)$$

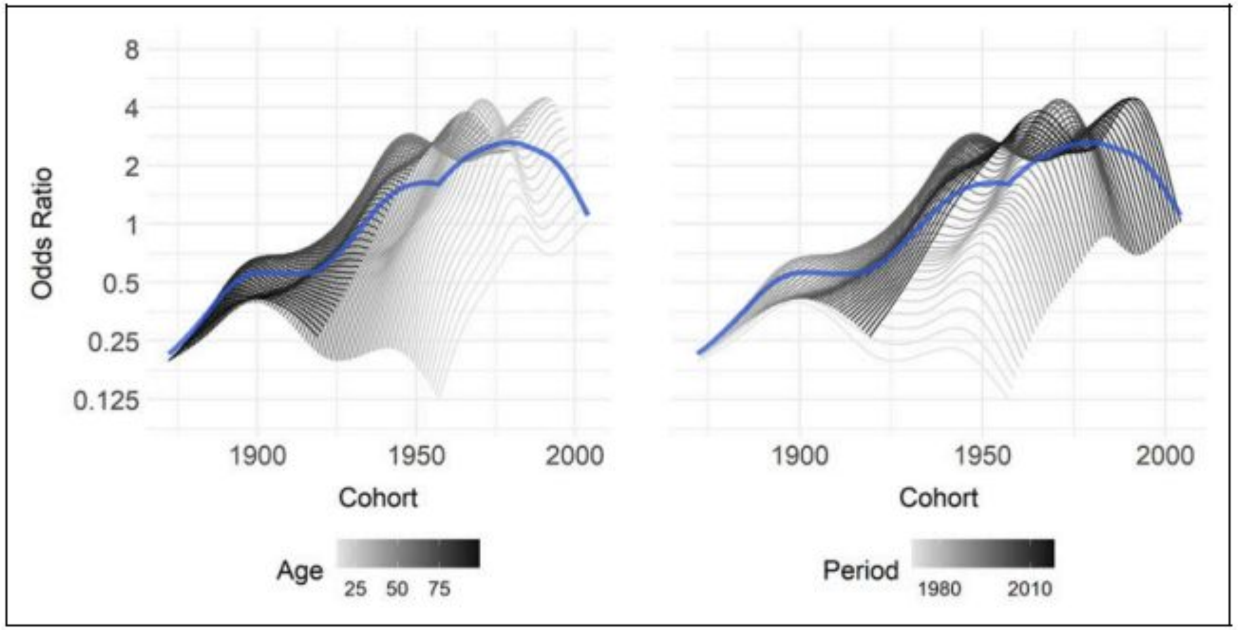
$$f_p\left(\text{period}_p\right) = \frac{1}{A}\sum_{a=1}^{A} f_{ap}\left(\text{age}_a \mid \text{period}_p\right)$$

$$f_c(\text{cohort}_c) = \frac{1}{A \times P}\sum_{a=1}^{A}\sum_{p=1}^{P} f_{ap}\left(\text{age}_a, \text{period}_p \mid \text{cohort}_c\right)$$

Examples - Weigert et al. (2021)

Table 3. Overview of marginal effects of the pure APC model (see Figure 7).

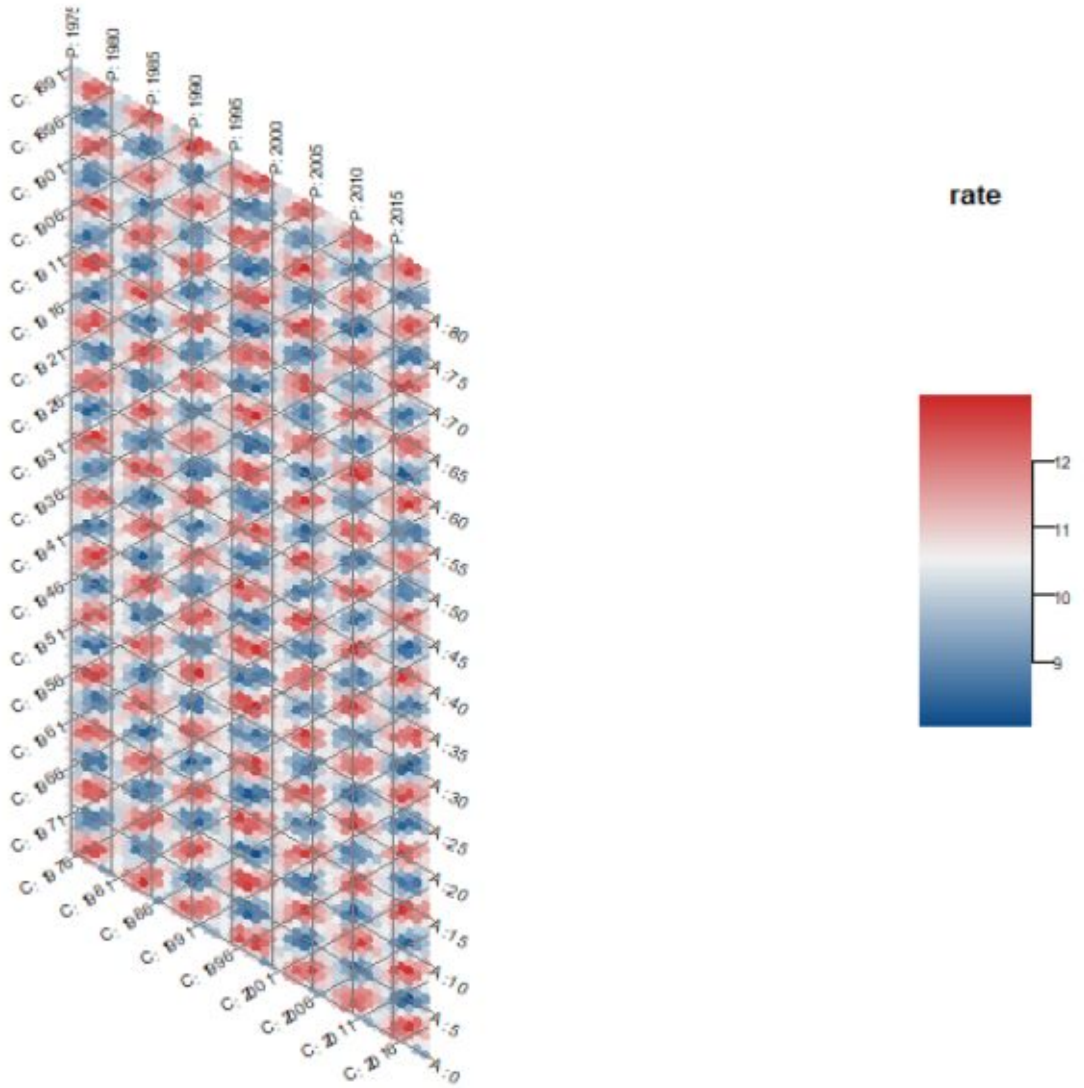
Model	Effect	Value with maximum OR	Value with minimum OR	Maximum OR	Minimum OR	Ratio
<500 km	Age	88	23	2.94	0.39	**7.5**
	Period	1971	2018	2.13	0.66	3.2
	Cohort	1939	1989	0.77	0.36	2.1
500–1000 km	Age	14	99	1.49	0.59	**2.5**
	Period	1983	2018	1.14	0.85	1.3
	Cohort	2004	1989	1.30	0.97	1.3
1000–2000 km	Age	22	86	2.27	0.39	**5.8**
	Period	2018	1971	1.51	0.46	3.3
	Cohort	1994	2004	2.50	1.35	1.9
2000–6000 km	Age	25	99	2.54	0.14	18.1
	Period	2009	1971	3.01	0.14	**21.5**
	Cohort	2004	1939	6.54	1.97	3.3
>6000 km	Age	27	99	2.10	0.32	6.6
	Period	2018	1971	1.75	0.25	**7.0**
	Cohort	1979	2004	2.64	1.11	2.4

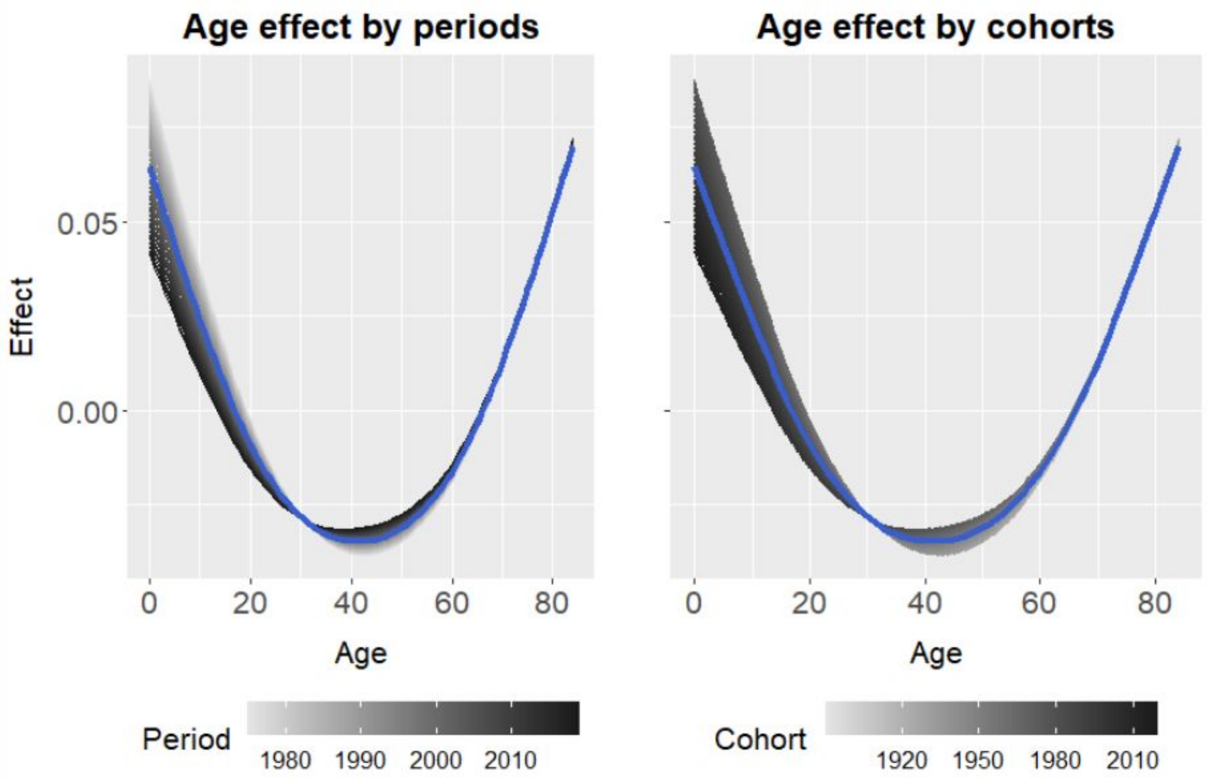


Examples

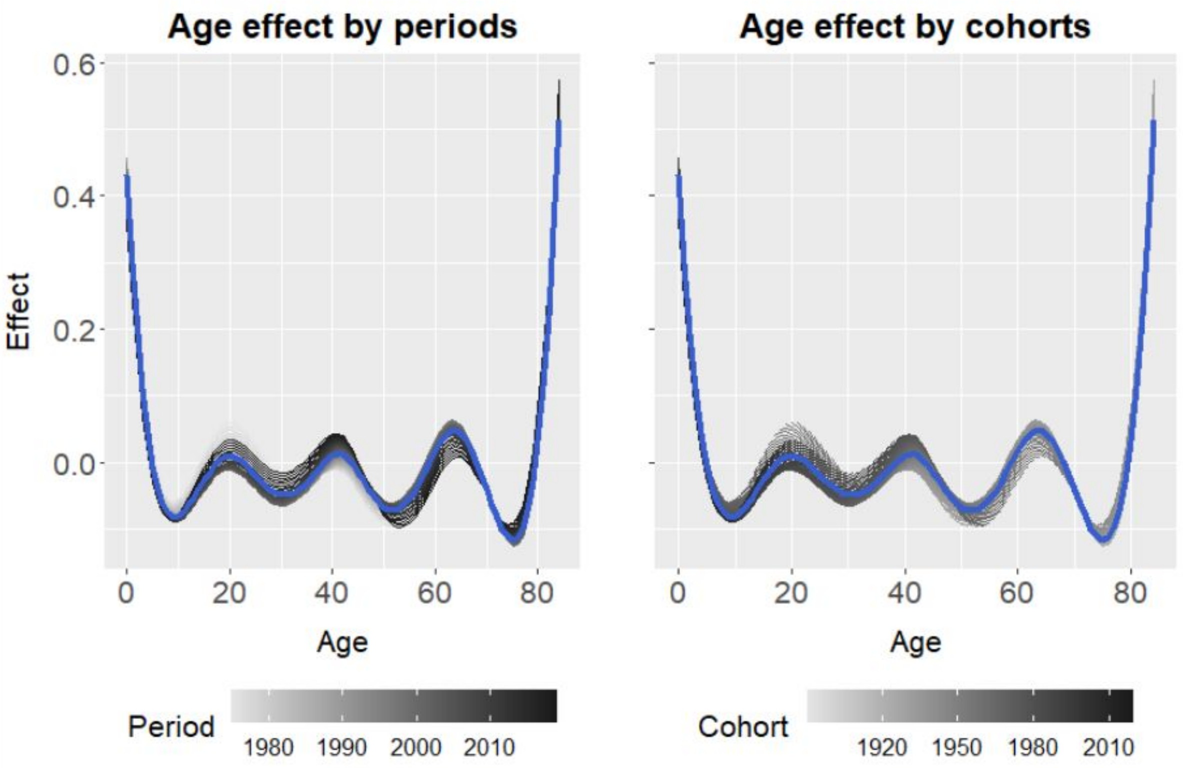
Using simulated data as example

- Sinusoidal, AC, noise level = 0.1
- Underlying data driven by Age and Cohort
- Comparing te(A,P) with te(A,P,C), (1 & 2 respectively)
- Partial Effect Plots

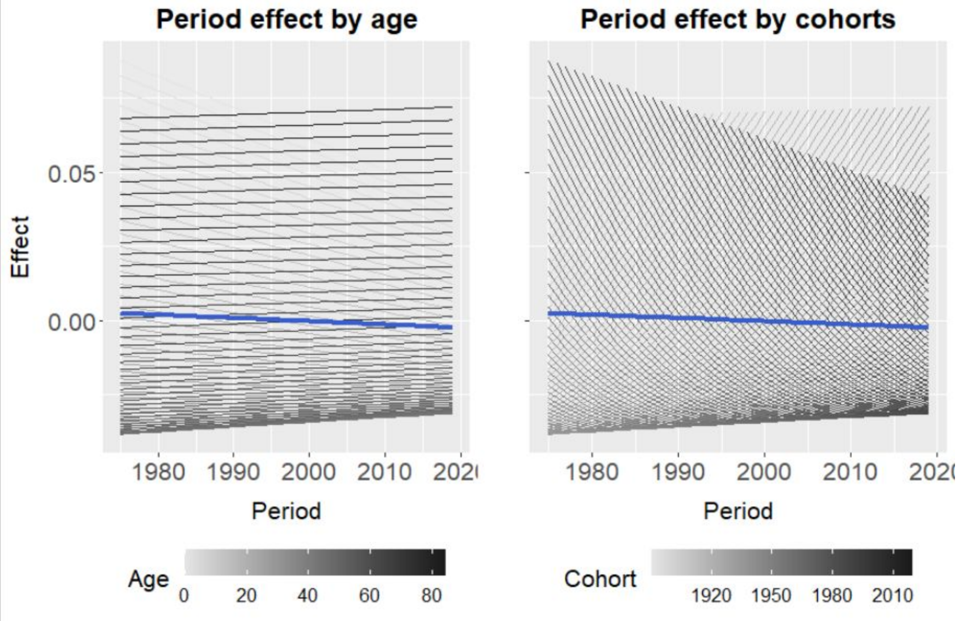




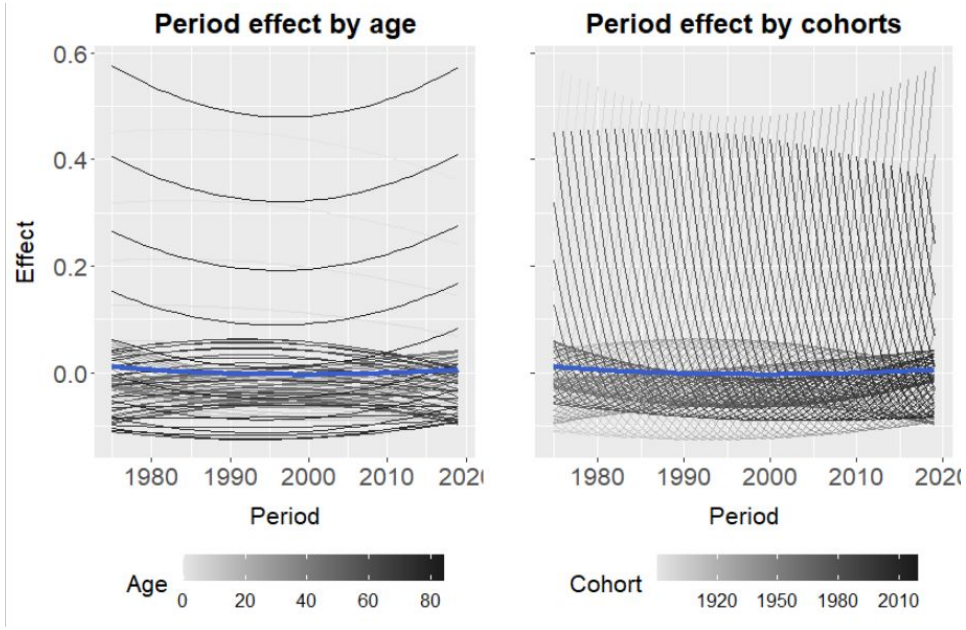
Model 1 te(A,P)



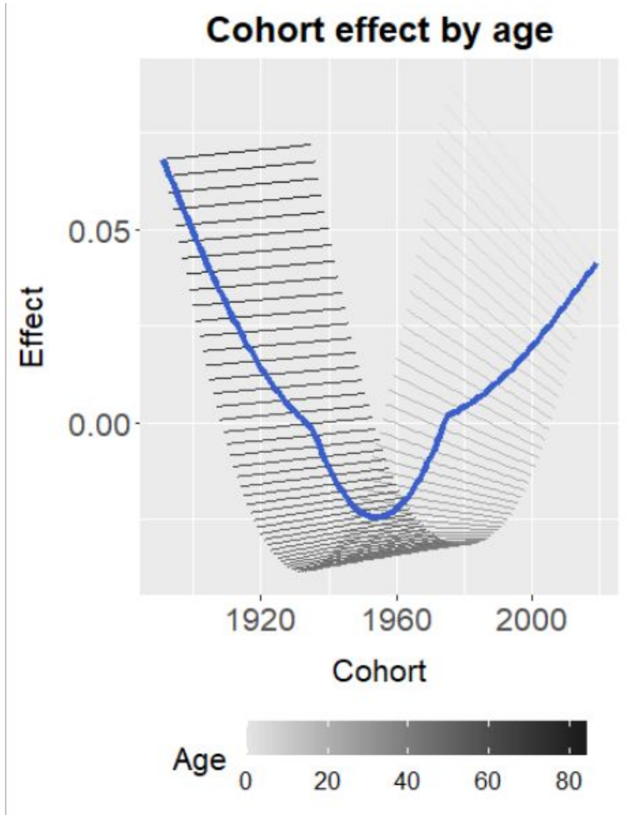
Model 2 te(A,P,C)



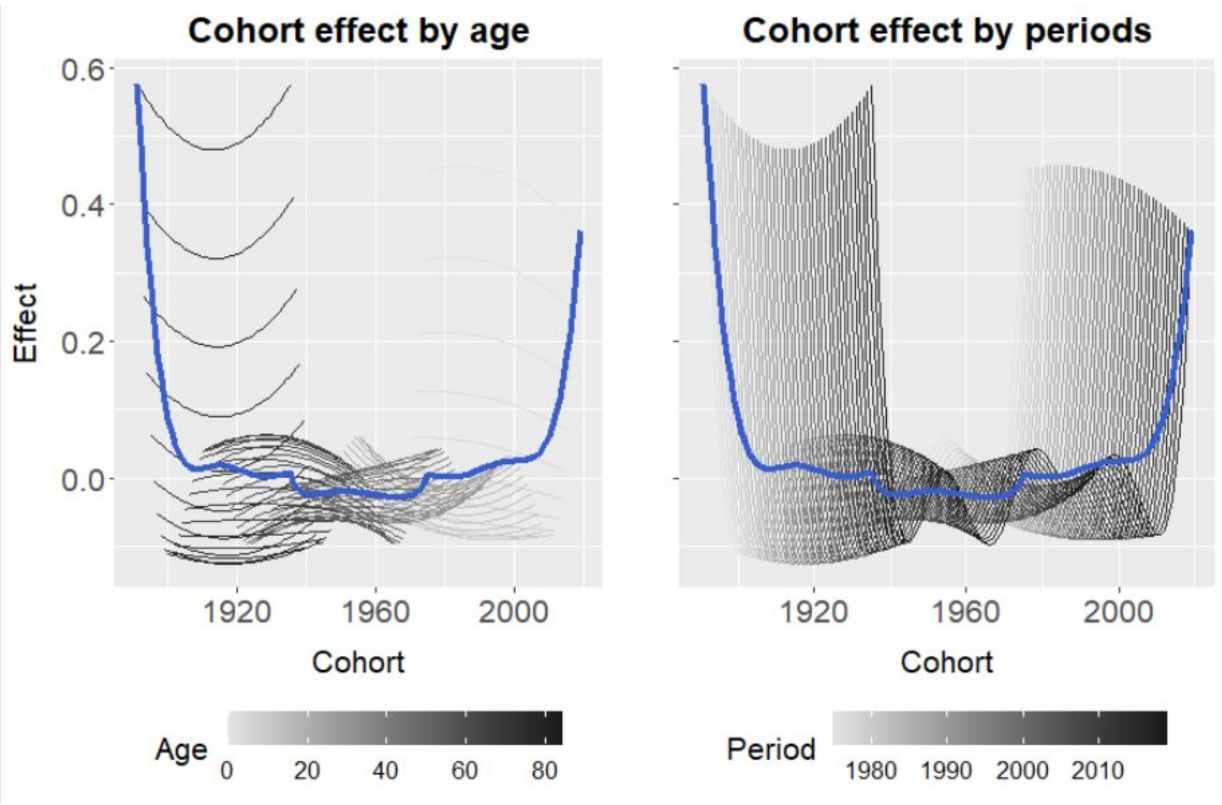
Model 1 te(A,P)



Model 2 te(A,P,C)



Model 1 te(A,P)



Model 2 te(A,P,C)

Examples

Using simulated data as example

- Sinusoidal, AC, noise level = 0.1
- Comparing te(A,P) with te(A,P,C), (1 & 2 respectively)
- Partial Effect Plots

model	effect	value_withMaxEffect	value_withMinEffect	max_effect	min_effect	difference
model 1	age	84	41	0.07	-0.03	0.10
model 1	period	1975	2019	0.00	0.00	0.00
model 1	cohort	1891	1954	0.07	-0.02	0.09
model 2	age	84	75	0.51	-0.12	0.63
model 2	period	1975	1999	0.01	0.00	0.01
model 2	cohort	1891	1967	0.58	-0.03	0.61