

## Summary Report

*Please review the whole document thoroughly before working on this report*

*Colors of Cells to be Filled by Tasks*

Task #1	Task 2	Task 3	Task 4
---------	--------	--------	--------

### Task #1

#### Variables

Role of Variable	Name of Variable in SAS codes	Type of Variable
Outcome Variable	HAVARTH4	Categorical (= 2 levels)
Explanatory Variable	BMI	Numeric
Explanatory Variable	PHYSHLTH	Numeric
Explanatory Variable	EXERANY2	Categorical (= 2 levels)
Explanatory Variable	SEX	Categorical (= 2 levels)
Explanatory Variable	AGEGROUP	Categorical (> 2 levels)
Explanatory Variable	SMOKDAY2	Categorical (> 2 levels)
Explanatory Variable of Interest (a modifiable risk factor, among those 6 explanatory variables listed above)	EXERANY2	

#### Article

Item	Values
PMID	36166875
Title	Association between work physical activity, dietary factors, and risk of rheumatoid arthritis
Table Legend (of the table that summarizes the results of variables which included in the simple or multiple logistic regression analysis)	Table 3. Association between leisure and work physical activity and incident rheumatoid arthritis (RA)

**Task #2, Task #3, and Task #4**

**Summary Table**

Explanatory Variable	Value N (%)	Outcome Variable: HAVARTH4		P-value of comparing groups	Odds Ratio [95% CI]	P-value	Adjusted Odds Ratio [95% CI]	Adjusted P-value
		Yes (N = 1265)	No (N = 758)					
BMI	Median [IQR]	29.29 [25.09, 34.33]	28.16 [24.41, 32.61]	0.0006	1.021 [1.009, 1.033]	0.0007	1.020 [1.007, 1.034]	0.0028
PHYSHLTH	Median [IQR]	7 [0, 25]	0 [0, 7]	<.0001	1.046 [1.037, 1.056]	<.0001	1.044 [1.035, 1.054]	<.0001
EXERANY2	Yes	694 (54.86)	518 (68.34)	<.0001	0.563 [0.466, 0.680]	<.0001	0.815 [0.663, 1.002]	0.0518
	No	571 (45.14)	240 (31.66)		1		1	
SEX	Male	412 (32.57)	345 (45.51)	<.0001	1	<.0001	1	
	Female	853 (67.43)	413 (54.49)		1.729 [1.437, 2.081]		1.818 [1.493, 2.214]	<.0001
AGEGROUP	Age 45 - 54	231 (18.26)	245 (32.32)	<.0001	1	*<.0001	1	
	Age 55 - 64	367 (29.01)	190 (25.07)		2.049 [1.594, 2.633]	<.0001	2.043 [1.569, 2.661]	<.0001
	Age 65 or older	667 (52.73)	323 (42.61)		2.190 [1.751, 2.739]	<.0001	2.654 [2.087, 3.376]	<.0001
SMOKDAY2	Every day	268 (21.19)	177 (23.35)	0.2385	0.900 [0.724, 1.120]	0.3453	N/A	
	Some days	94 (7.43)	44 (5.80)		1.270 [0.874, 1.846]	0.2092		
	Not at all	903 (71.38)	537 (70.84)		1	*0.2398		

## Task #4

### Model Selection

<b>Model Selection Approach for the Final Model</b>	Backward-elimination
---	----------------------

(One of the possible values: Backward-elimination / Forward selection / Stepwise selection)

### Names of 3 Variables which will be remained in the model selection process

(need to have all 3 types of variables: Numeric, Categorical (= 2 levels), Categorical (>2 levels))

Role	Name of Variable	Type of Variable
Explanatory Variable of Interest (a modifiable risk factor)	EXERANY2	Categorical (=2 levels)
Explanatory Variable	BMI	Numeric
Explanatory Variable	AGEGROUP	Categorical (>2 levels)

### Results from Model Selection: Names of Variables based on the Model Selection Approach

(Use "N/A" to fill the following table if no associated variables are reported in the SAS output)

Backward-elimination	Forward	Stepwise	
(under "Effect Removed")	(under "Effect Entered")	(under "Effect, Entered")	(under "Effect, Removed")
SMOKDAY2	PHYSHLTH	PHYSHLTH	N/A
N/A	SEX	SEX	N/A
N/A	N/A	N/A	N/A

### Results of Model Fitting

Test Statistics	Value
R-square	0.1049
Max-rescaled R-Square	0.1430
Hosmer and Lemeshow Goodness-of-Fit Test ("Chi-Square")	5.1759
Hosmer and Lemeshow Goodness-of-Fit Test ("Pr > ChiSq")	0.7386
Area Under the Curve	0.6927