[1]-- The extraction of comorbidities is based on the official visualization view "charlson" of the MIMIC database, the code is as follows.

------------------------------------------------------------------

-- This query extracts Charlson Comorbidity Index (CCI) based on the

-- recorded ICD-9 and ICD-10 codes.

--

-- Reference for CCI:

-- (1) Charlson ME, Pompei P, Ales KL, MacKenzie CR. (1987) A new method

-- of classifying prognostic comorbidity in longitudinal studies:

-- development and validation.J Chronic Dis; 40(5):373-83.

--

-- (2) Charlson M, Szatrowski TP, Peterson J, Gold J. (1994) Validation

-- of a combined comorbidity index. J Clin Epidemiol; 47(11):1245-51.

--

-- Reference for ICD-9-CM and ICD-10 Coding Algorithms for Charlson

-- Comorbidities:

-- (3) Quan H, Sundararajan V, Halfon P, et al. Coding algorithms for

-- defining Comorbidities in ICD-9-CM and ICD-10 administrative data.

-- Med Care. 2005 Nov; 43(11): 1130-9.

-- ------------------------------------------------------------------

WITH diag AS (

SELECT

hadm\_id

, CASE WHEN icd\_version = 9 THEN icd\_code ELSE NULL END AS icd9\_code

, CASE WHEN icd\_version = 10 THEN icd\_code ELSE NULL END AS icd10\_code

FROM `physionet-data.mimiciv\_hosp.diagnoses\_icd`

)

, com AS (

SELECT

ad.hadm\_id

-- Myocardial infarction

, MAX(CASE WHEN

SUBSTR(icd9\_code, 1, 3) IN ('410', '412')

OR

SUBSTR(icd10\_code, 1, 3) IN ('I21', 'I22')

OR

SUBSTR(icd10\_code, 1, 4) = 'I252'

THEN 1

ELSE 0 END) AS myocardial\_infarct

-- Congestive heart failure

, MAX(CASE WHEN

SUBSTR(icd9\_code, 1, 3) = '428'

OR

SUBSTR(

icd9\_code, 1, 5

) IN ('39891', '40201', '40211', '40291', '40401', '40403'

, '40411', '40413', '40491', '40493')

OR

SUBSTR(icd9\_code, 1, 4) BETWEEN '4254' AND '4259'

OR

SUBSTR(icd10\_code, 1, 3) IN ('I43', 'I50')

OR

SUBSTR(

icd10\_code, 1, 4

) IN ('I099', 'I110', 'I130', 'I132', 'I255', 'I420'

, 'I425', 'I426', 'I427', 'I428', 'I429', 'P290'

)

THEN 1

ELSE 0 END) AS congestive\_heart\_failure

-- Peripheral vascular disease

, MAX(CASE WHEN

SUBSTR(icd9\_code, 1, 3) IN ('440', '441')

OR

SUBSTR(

icd9\_code, 1, 4

) IN ('0930', '4373', '4471', '5571', '5579', 'V434')

OR

SUBSTR(icd9\_code, 1, 4) BETWEEN '4431' AND '4439'

OR

SUBSTR(icd10\_code, 1, 3) IN ('I70', 'I71')

OR

SUBSTR(icd10\_code, 1, 4) IN ('I731', 'I738', 'I739', 'I771', 'I790'

, 'I792'

, 'K551'

, 'K558'

, 'K559'

, 'Z958'

, 'Z959'

)

THEN 1

ELSE 0 END) AS peripheral\_vascular\_disease

-- Cerebrovascular disease

, MAX(CASE WHEN

SUBSTR(icd9\_code, 1, 3) BETWEEN '430' AND '438'

OR

SUBSTR(icd9\_code, 1, 5) = '36234'

OR

SUBSTR(icd10\_code, 1, 3) IN ('G45', 'G46')

OR

SUBSTR(icd10\_code, 1, 3) BETWEEN 'I60' AND 'I69'

OR

SUBSTR(icd10\_code, 1, 4) = 'H340'

THEN 1

ELSE 0 END) AS cerebrovascular\_disease

-- Dementia

, MAX(CASE WHEN

SUBSTR(icd9\_code, 1, 3) = '290'

OR

SUBSTR(icd9\_code, 1, 4) IN ('2941', '3312')

OR

SUBSTR(icd10\_code, 1, 3) IN ('F00', 'F01', 'F02', 'F03', 'G30')

OR

SUBSTR(icd10\_code, 1, 4) IN ('F051', 'G311')

THEN 1

ELSE 0 END) AS dementia

-- Chronic pulmonary disease

, MAX(CASE WHEN

SUBSTR(icd9\_code, 1, 3) BETWEEN '490' AND '505'

OR

SUBSTR(icd9\_code, 1, 4) IN ('4168', '4169', '5064', '5081', '5088')

OR

SUBSTR(icd10\_code, 1, 3) BETWEEN 'J40' AND 'J47'

OR

SUBSTR(icd10\_code, 1, 3) BETWEEN 'J60' AND 'J67'

OR

SUBSTR(icd10\_code, 1, 4) IN ('I278', 'I279', 'J684', 'J701', 'J703')

THEN 1

ELSE 0 END) AS chronic\_pulmonary\_disease

-- Rheumatic disease

, MAX(CASE WHEN

SUBSTR(icd9\_code, 1, 3) = '725'

OR

SUBSTR(icd9\_code, 1, 4) IN ('4465', '7100', '7101', '7102', '7103'

, '7104', '7140', '7141', '7142', '7148'

)

OR

SUBSTR(icd10\_code, 1, 3) IN ('M05', 'M06', 'M32', 'M33', 'M34')

OR

SUBSTR(icd10\_code, 1, 4) IN ('M315', 'M351', 'M353', 'M360')

THEN 1

ELSE 0 END) AS rheumatic\_disease

-- Peptic ulcer disease

, MAX(CASE WHEN

SUBSTR(icd9\_code, 1, 3) IN ('531', '532', '533', '534')

OR

SUBSTR(icd10\_code, 1, 3) IN ('K25', 'K26', 'K27', 'K28')

THEN 1

ELSE 0 END) AS peptic\_ulcer\_disease

-- Mild liver disease

, MAX(CASE WHEN

SUBSTR(icd9\_code, 1, 3) IN ('570', '571')

OR

SUBSTR(

icd9\_code, 1, 4

) IN ('0706', '0709', '5733', '5734', '5738', '5739', 'V427')

OR

SUBSTR(

icd9\_code, 1, 5

) IN ('07022', '07023', '07032', '07033', '07044', '07054')

OR

SUBSTR(icd10\_code, 1, 3) IN ('B18', 'K73', 'K74')

OR

SUBSTR(

icd10\_code, 1, 4

) IN ('K700', 'K701', 'K702', 'K703', 'K709', 'K713'

, 'K714', 'K715', 'K717', 'K760', 'K762'

, 'K763', 'K764', 'K768', 'K769', 'Z944')

THEN 1

ELSE 0 END) AS mild\_liver\_disease

-- Diabetes without chronic complication

, MAX(CASE WHEN

SUBSTR(

icd9\_code, 1, 4

) IN ('2500', '2501', '2502', '2503', '2508', '2509')

OR

SUBSTR(

icd10\_code, 1, 4

) IN ('E100', 'E101', 'E106', 'E108', 'E109', 'E110', 'E111'

, 'E116'

, 'E118'

, 'E119'

, 'E120'

, 'E121'

, 'E126'

, 'E128'

, 'E129'

, 'E130'

, 'E131'

, 'E136'

, 'E138'

, 'E139'

, 'E140'

, 'E141', 'E146', 'E148', 'E149')

THEN 1

ELSE 0 END) AS diabetes\_without\_cc

-- Diabetes with chronic complication

, MAX(CASE WHEN

SUBSTR(icd9\_code, 1, 4) IN ('2504', '2505', '2506', '2507')

OR

SUBSTR(

icd10\_code, 1, 4

) IN ('E102', 'E103', 'E104', 'E105', 'E107', 'E112', 'E113'

, 'E114'

, 'E115'

, 'E117'

, 'E122'

, 'E123'

, 'E124'

, 'E125'

, 'E127'

, 'E132'

, 'E133'

, 'E134'

, 'E135'

, 'E137'

, 'E142'

, 'E143', 'E144', 'E145', 'E147')

THEN 1

ELSE 0 END) AS diabetes\_with\_cc

-- Hemiplegia or paraplegia

, MAX(CASE WHEN

SUBSTR(icd9\_code, 1, 3) IN ('342', '343')

OR

SUBSTR(icd9\_code, 1, 4) IN ('3341', '3440', '3441', '3442'

, '3443', '3444', '3445', '3446', '3449'

)

OR

SUBSTR(icd10\_code, 1, 3) IN ('G81', 'G82')

OR

SUBSTR(icd10\_code, 1, 4) IN ('G041', 'G114', 'G801', 'G802', 'G830'

, 'G831'

, 'G832'

, 'G833'

, 'G834'

, 'G839'

)

THEN 1

ELSE 0 END) AS paraplegia

-- Renal disease

, MAX(CASE WHEN

SUBSTR(icd9\_code, 1, 3) IN ('582', '585', '586', 'V56')

OR

SUBSTR(icd9\_code, 1, 4) IN ('5880', 'V420', 'V451')

OR

SUBSTR(icd9\_code, 1, 4) BETWEEN '5830' AND '5837'

OR

SUBSTR(

icd9\_code, 1, 5

) IN (

'40301'

, '40311'

, '40391'

, '40402'

, '40403'

, '40412'

, '40413'

, '40492'

, '40493'

)

OR

SUBSTR(icd10\_code, 1, 3) IN ('N18', 'N19')

OR

SUBSTR(icd10\_code, 1, 4) IN ('I120', 'I131', 'N032', 'N033', 'N034'

, 'N035'

, 'N036'

, 'N037'

, 'N052'

, 'N053'

, 'N054'

, 'N055'

, 'N056'

, 'N057'

, 'N250'

, 'Z490'

, 'Z491'

, 'Z492'

, 'Z940'

, 'Z992'

)

THEN 1

ELSE 0 END) AS renal\_disease

-- Any malignancy, including lymphoma and leukemia,

-- except malignant neoplasm of skin.

, MAX(CASE WHEN

SUBSTR(icd9\_code, 1, 3) BETWEEN '140' AND '172'

OR

SUBSTR(icd9\_code, 1, 4) BETWEEN '1740' AND '1958'

OR

SUBSTR(icd9\_code, 1, 3) BETWEEN '200' AND '208'

OR

SUBSTR(icd9\_code, 1, 4) = '2386'

OR

SUBSTR(icd10\_code, 1, 3) IN ('C43', 'C88')

OR

SUBSTR(icd10\_code, 1, 3) BETWEEN 'C00' AND 'C26'

OR

SUBSTR(icd10\_code, 1, 3) BETWEEN 'C30' AND 'C34'

OR

SUBSTR(icd10\_code, 1, 3) BETWEEN 'C37' AND 'C41'

OR

SUBSTR(icd10\_code, 1, 3) BETWEEN 'C45' AND 'C58'

OR

SUBSTR(icd10\_code, 1, 3) BETWEEN 'C60' AND 'C76'

OR

SUBSTR(icd10\_code, 1, 3) BETWEEN 'C81' AND 'C85'

OR

SUBSTR(icd10\_code, 1, 3) BETWEEN 'C90' AND 'C97'

THEN 1

ELSE 0 END) AS malignant\_cancer

-- Moderate or severe liver disease

, MAX(CASE WHEN

SUBSTR(icd9\_code, 1, 4) IN ('4560', '4561', '4562')

OR

SUBSTR(icd9\_code, 1, 4) BETWEEN '5722' AND '5728'

OR

SUBSTR(

icd10\_code, 1, 4

) IN ('I850', 'I859', 'I864', 'I982', 'K704', 'K711'

, 'K721', 'K729', 'K765', 'K766', 'K767')

THEN 1

ELSE 0 END) AS severe\_liver\_disease

-- Metastatic solid tumor

, MAX(CASE WHEN

SUBSTR(icd9\_code, 1, 3) IN ('196', '197', '198', '199')

OR

SUBSTR(icd10\_code, 1, 3) IN ('C77', 'C78', 'C79', 'C80')

THEN 1

ELSE 0 END) AS metastatic\_solid\_tumor

-- AIDS/HIV

, MAX(CASE WHEN

SUBSTR(icd9\_code, 1, 3) IN ('042', '043', '044')

OR

SUBSTR(icd10\_code, 1, 3) IN ('B20', 'B21', 'B22', 'B24')

THEN 1

ELSE 0 END) AS aids

FROM `physionet-data.mimiciv\_hosp.admissions` ad

LEFT JOIN diag

ON ad.hadm\_id = diag.hadm\_id

GROUP BY ad.hadm\_id

)

, ag AS (

SELECT

hadm\_id

, age

, CASE WHEN age <= 50 THEN 0

WHEN age <= 60 THEN 1

WHEN age <= 70 THEN 2

WHEN age <= 80 THEN 3

ELSE 4 END AS age\_score

FROM `physionet-data.mimiciv\_derived.age`

)

SELECT

ad.subject\_id

, ad.hadm\_id

, ag.age\_score

, myocardial\_infarct

, congestive\_heart\_failure

, peripheral\_vascular\_disease

, cerebrovascular\_disease

, dementia

, chronic\_pulmonary\_disease

, rheumatic\_disease

, peptic\_ulcer\_disease

, mild\_liver\_disease

, diabetes\_without\_cc

, diabetes\_with\_cc

, paraplegia

, renal\_disease

, malignant\_cancer

, severe\_liver\_disease

, metastatic\_solid\_tumor

, aids

-- Calculate the Charlson Comorbidity Score using the original

-- weights from Charlson, 1987.

, age\_score

+ myocardial\_infarct + congestive\_heart\_failure

+ peripheral\_vascular\_disease + cerebrovascular\_disease

+ dementia + chronic\_pulmonary\_disease

+ rheumatic\_disease + peptic\_ulcer\_disease

+ GREATEST(mild\_liver\_disease, 3 \* severe\_liver\_disease)

+ GREATEST(2 \* diabetes\_with\_cc, diabetes\_without\_cc)

+ GREATEST(2 \* malignant\_cancer, 6 \* metastatic\_solid\_tumor)

+ 2 \* paraplegia + 2 \* renal\_disease

+ 6 \* aids

AS charlson\_comorbidity\_index

FROM `physionet-data.mimiciv\_hosp.admissions` ad

LEFT JOIN com

ON ad.hadm\_id = com.hadm\_id

LEFT JOIN ag

ON com.hadm\_id = ag.hadm\_id

;

**References:**

[1]. Quan, H., et al., Coding algorithms for defining comorbidities in ICD-9-CM and ICD-10 administrative data. Med Care, 2005. 43(11): p. 1130-9.