

```
LIBNAME survey XLSX "/home/u63581499/SURVEYDATA.xlsx";

PROC IMPORT DATAFILE="/home/u63581499/SURVEYDATA.xlsx"
OUT=survey_data
DSM5=EXCEL
REPLACE;
SHEET="SHEET1";
GETNAMES=YES;
RUN;

proc freq DATA=survey_data;
Tables Household_income;
run;

proc freq DATA=survey_data;
Tables Household_income /MISSING;
run;

/* RECODEING***** */
DATA survey_data_recode;
SET survey_data;

/****** Age *****/
/* Keep only 18-45 years */
IF NOT (18 <= Age <= 45) THEN DELETE;

/* Continuous age recode to 18-45 */
age1 = Age;

/* Binary age: 0=18-26; 1=27-45 */
IF 18 <= Age <= 26 THEN age2 = 0;
ELSE IF 27 <= Age <= 45 THEN age2 = 1;
ELSE age2 = .;

/****** Gender *****/
IF STRIP(UPCASE(Gender)) = "MALE" THEN gender1 = "Male";
ELSE IF STRIP(UPCASE(Gender)) = "FEMALE" THEN gender1 = "Female";
ELSE IF UPCASE(Gender) = "PREFER NOT TO SAY" THEN gender1 = "Prefer not to say";
ELSE gender1 = "Other";

/****** Foreign-born *****/
IF UPCASE(STRIP(Born)) = "YES" THEN foreign_born = "Yes";
ELSE IF UPCASE(STRIP(Born)) = "NO" THEN foreign_born = "No";
ELSE foreign_born = "Prefer not to say";

/****** Race / Ethnicity *****/
IF UPCASE(Hispanic_spanish_latino) = "NO" THEN DO;
IF UPCASE(Race) = "WHITE" THEN racel = "Non-Hispanic White";
ELSE IF UPCASE(Race) = "BLACK OR AFRICAN AMERICAN" THEN racel = "Non-Hispanic Black";
ELSE IF UPCASE(Race) IN ("ASIAN","SOUTH ASIAN",
THEN racel = "Asian";
ELSE racel = "Other";
ELSE IF UPCASE(Hispanic_spanish_latino) = "YES" THEN racel = "Hispanic";
ELSE IF UPCASE(Hispanic_spanish_latino) = "PREFER NOT TO SAY" OR UPCASE(Race) = "PREFER NOT TO SAY" THEN racel = "Prefer not to say";
ELSE racel = "Other";

/****** Household income *****/
SELECT (STRIP(Household_income));
WHEN ('$21-39,000', 'Less than $20,000') income = '<39,000';
WHEN ('$40-59,000', '$60-79,000') income = '40-79,000';
WHEN ('$80-99,000', '$100-149,000', '$150,000 or more') income = '80,000+';
WHEN ('Prefer not to say', 'Unsure') income = 'Unsure';
OTHERWISE income = 'Unsure'; /* catch-all */
END;

/****** Education *****/
IF Education IN ("Less than 12th grade", "High school (completed 12th grade)") THEN education1 = "High school or less";
ELSE IF Education IN ("Some college","Associates","Bachelors") THEN education1 = "Bachelors or less of college";
ELSE IF Education IN ("Masters","Doctorate") THEN education1 = "Graduate degree";
ELSE education1 = "Other";

/****** Likert scale recode vac questions: safety, importance, effectiveness, belief *****/
/* 1 = strongly agree (pro-vaccine), 5 = strongly disagree */

SELECT (UPCASE(STRIP(Cu_Fpvax_safe))));
WHEN ("STRONGLY AGREE") safe_score = 1;
WHEN ("SOMEWHAT AGREE") safe_score = 2;
WHEN ("DO NOT KNOW") safe_score = 3;
WHEN ("SOMEWHAT DISAGREE") safe_score = 4;
WHEN ("STRONGLY DISAGREE") safe_score = 5;
OTHERWISE safe_score = .;
END;

SELECT (UPCASE(STRIP(Cu_Fpvax_imp))));
WHEN ("STRONGLY AGREE") imp_score = 1;
WHEN ("SOMEWHAT AGREE") imp_score = 2;
WHEN ("DO NOT KNOW") imp_score = 3;
WHEN ("SOMEWHAT DISAGREE") imp_score = 4;
WHEN ("STRONGLY DISAGREE") imp_score = 5;
OTHERWISE imp_score = .;
END;

SELECT (UPCASE(STRIP(Cu_Fpvax_eff))));
WHEN ("STRONGLY AGREE") eff_score = 1;
WHEN ("SOMEWHAT AGREE") eff_score = 2;
WHEN ("DO NOT KNOW") eff_score = 3;
WHEN ("SOMEWHAT DISAGREE") eff_score = 4;
WHEN ("STRONGLY DISAGREE") eff_score = 5;
OTHERWISE eff_score = .;
END;

SELECT (UPCASE(STRIP(Cu_Fpvax_b))));
WHEN ("STRONGLY AGREE") belief_score = 1;
WHEN ("SOMEWHAT AGREE") belief_score = 2;
WHEN ("DO NOT KNOW") belief_score = 3;
WHEN ("SOMEWHAT DISAGREE") belief_score = 4;
WHEN ("STRONGLY DISAGREE") belief_score = 5;
OTHERWISE belief_score = .;
END;

/* Continuous scores: higher = more negative attitudes / more hesitancy */
score = SUM(safe_score, imp_score, eff_score, belief_score);
score2 = SUM(safe_score, imp_score, eff_score);

/* Binary scoreb with cutoff= 5 :
score <=5 = lower hesitancy (0); score >5 = higher hesitancy (1) */
IF score NE . THEN DO;
IF score > 5 THEN scoreb = 1;
ELSE scoreb = 0;
END;
ELSE scoreb = .;

/****** Binary versions of vac questions *****/
IF UPCASE(Cu_Fpvax_imp) IN ("STRONGLY AGREE","SOMEWHAT AGREE") THEN Vaccine_Important = "Agree";
ELSE Vaccine_Important = "Disagree";

IF UPCASE(Cu_Fpvax_safe) IN ("STRONGLY AGREE","SOMEWHAT AGREE") THEN Vaccine_Safe = "Agree";
ELSE Vaccine_Safe = "Disagree";

IF UPCASE(Cu_Fpvax_eff) IN ("STRONGLY AGREE","SOMEWHAT AGREE") THEN Vaccine_Effective = "Agree";
ELSE Vaccine_Effective = "Disagree";

IF UPCASE(Cu_Fpvax_b) IN ("STRONGLY AGREE","SOMEWHAT AGREE") THEN Personal_Belief_Compatible = "Agree";
ELSE Personal_Belief_Compatible = "Disagree";

/****** Hesitancy/refusal reason *****/
IF UPCASE(Vaccine_refusal) = "YES" OR UPCASE(Vaccine_hesitancy) = "YES" THEN reason = 1;
ELSE reason = 0;

/****** Reasons *****/
LENGTH
did_not_think_needed did_not_think_effective did_not_think_safe
not_enough_data bad_experience_vaccine bad_experience_clinic
provider_restriction other_beliefs religious_reasons
work_timing_inconvenient too_far_no_reason;

IF reason = 1 THEN DO;
LENGTH reason_txt $1000;
reason_txt = UPCASE(Vaccine_refusal_reason);

did_not_think_needed = (INDEX(reason_txt,
"DID NOT THINK IT WAS NEEDED") > 0);
did_not_think_effective = (INDEX(reason_txt,
"DID NOT THINK IT WAS EFFECTIVE") > 0);
did_not_think_safe = (INDEX(reason_txt,
"DID NOT THINK IT WAS SAFE") > 0);
not_enough_data = (INDEX(reason_txt,
"NOT ENOUGH DATA/INFORMATION RELATED TO MY DEMOGRAPHIC") > 0);
bad_experience_vaccine = (INDEX(reason_txt,
"HAD A BAD EXPERIENCE OR REACTION WITH PREVIOUS VACCINATION") > 0);
bad_experience_clinic = (INDEX(reason_txt,
"HAD A BAD EXPERIENCE WITH PREVIOUS VACCINATOR/HEALTH CLINIC") > 0);
provider_restriction = (INDEX(reason_txt,
"HEALTHCARE PROVIDER TOLD ME I CANNOT TAKE IT DUE TO UNDERLYING HEALTH PROBLEM OR ALLERGY") > 0);
other_beliefs = (INDEX(reason_txt,
"OTHER BELIEFS/RELY ON TRADITIONAL MEDICINE") > 0);
religious_reasons = (INDEX(reason_txt,
"RELIGIOUS REASONS") > 0);
work_conflict = (INDEX(reason_txt,
"NOT POSSIBLE TO LEAVE OTHER WORK (AT HOME OR JOB)") > 0);
timing_inconvenient = (INDEX(reason_txt,
"TIMING INCONVENIENT") > 0);
work_timing_inconvenient = MAX(work_conflict, timing_inconvenient);
too_far = (INDEX(reason_txt,
"TOO FAR AWAY TO GET VACCINE") > 0);
no_reason = (INDEX(reason_txt,
"DO NOT KNOW/CANNOT REMEMBER/NO REASON") > 0);
END;
ELSE DO;
did_not_think_needed = .;
did_not_think_effective = .;
did_not_think_safe = .;
not_enough_data = .;
bad_experience_vaccine = .;
bad_experience_clinic = .;
provider_restriction = .;
other_beliefs = .;
religious_reasons = .;
work_conflict = .;
timing_inconvenient = .;
work_timing_inconvenient = .;
too_far = .;
no_reason = .;
END;

/****** HPV vaccination status *****/
IF INDEX(UPCASE(adult_vax), "HPV") > 0 THEN hpv = "Yes";
ELSE hpv = "No";
RUN;

/* ***** FINAL ANALYSIS*/
DATA survey_final;
SET survey_data_recode;
KEEP age1 age2 gender1 foreign_born racel income education1
Vaccine_hesitancy Vaccine_refusal
Vaccine_Important Vaccine_Safe Vaccine_Effective
Personal_Belief_Compatible hpv;
/****** DESCRIPTIVES*****/

/* Score distribution */
PROC MEANS DATA=survey_data_recode N MEAN MEDIAN MIN Q1 Q3 MAX;
VAR score;
RUN;

/* Basic descriptives */
PROC FREQ DATA=survey_final;
TABLES age2 gender1 foreign_born racel income education1
vaccine_hesitancy vaccine_refusal
vaccine_important vaccine_safe vaccine_effective
Personal_Belief_Compatible hpv / MISSING;
RUN;

PROC MEANS DATA=survey_final;
VAR age1;
RUN;

/* Reason counts */
PROC MEAS DATA=survey_data_recode SUM NOPRINT;
WHERE UPCASE(Vaccine_refusal) = "YES";
VAR did_not_think_needed did_not_think_effective did_not_think_safe
not_enough_data bad_experience_vaccine bad_experience_clinic
provider_restriction other_beliefs religious_reasons
work_timing_inconvenient too_far_no_reason;
OUTPUT OUT=reason_counts_ref SUM;
RUN;

PROC MEANS DATA=survey_data_recode SUM NOPRINT;
WHERE UPCASE(Vaccine_hesitancy) = "YES";
VAR did_not_think_needed did_not_think_effective did_not_think_safe
not_enough_data bad_experience_vaccine bad_experience_clinic
provider_restriction other_beliefs religious_reasons
work_timing_inconvenient too_far_no_reason;
OUTPUT OUT=reason_counts_hes SUM;
RUN;

/* check */
PROC PRINT DATA=survey_final (OBS=10);
RUN;

/****** TABLE 1: AGE & DEMOGRAPHICS x HPV STATUS*****/
PROC MEANS DATA=survey_final N MEAN STDDEV;
VAR age1;
RUN;

PROC MEANS DATA=survey_final N MEAN STDDEV;
CLASS hpv;
VAR age1;
RUN;

PROC TTEST DATA=survey_final;
CLASS hpv;
VAR age1;
RUN;

PROC FREQ DATA=survey_final;
TABLES age2 gender1 foreign_born racel income education1 hpv / MISSING;
RUN;

PROC FREQ DATA=survey_final;
TABLES
age2 * hpv
gender1 * hpv
foreign_born * hpv
racel * hpv
income * hpv
education1 * hpv
/ CHISQ NOROW NOPERCENT MISSING;
RUN;

/****** TABLE 2: VACCINE ATTITUDES x HPV***** */
PROC FREQ DATA=survey_final;
TABLES Vaccine_Important Vaccine_Safe Vaccine_Effective
Personal_Belief_Compatible
Vaccine_hesitancy Vaccine_refusal / MISSING;
RUN;

PROC FREQ DATA=survey_final;
TABLES
Vaccine_Important * hpv
Vaccine_Safe * hpv
Vaccine_Effective * hpv
Personal_Belief_Compatible * hpv
Vaccine_hesitancy * hpv
Vaccine_refusal * hpv
/ CHISQ NOROW MISSING;
RUN;

/****** REASONING VARIABLES x HPV***** */
PROC SQL;
SELECT
COUNT(*) AS Total_Respondents,
SUM(reason) AS Respondents_Yes,
100 * SUM(reason) / COUNT(*) AS Percentage_Yes,
COUNT(*) AS Respondents_No,
100 * (COUNT(*) - SUM(reason)) / COUNT(*) AS Percentage_No
FROM survey_data_recode;
QUIT;

PROC FREQ DATA=survey_data_recode;
TABLES reason * hpv / CHISQ NOROW NOPERCENT MISSING;
RUN;

/* Ever hesitant or refused */
PROC FREQ DATA=survey_data_recode;
WHERE UPCASE(Vaccine_hesitancy) = "YES"
OR UPCASE(Vaccine_refusal) = "YES";
TABLES did_not_think_needed did_not_think_effective did_not_think_safe
not_enough_data bad_experience_vaccine bad_experience_clinic
provider_restriction other_beliefs religious_reasons
work_timing_inconvenient too_far_no_reason / NOROW MISSING;
RUN;

PROC FREQ DATA=survey_data_recode;
WHERE UPCASE(Vaccine_hesitancy) = "YES"
OR UPCASE(Vaccine_refusal) = "YES";
TABLES
did_not_think_needed * hpv
did_not_think_effective * hpv
did_not_think_safe * hpv
not_enough_data * hpv
bad_experience_vaccine * hpv
bad_experience_clinic * hpv
provider_restriction * hpv
other_beliefs * hpv
religious_reasons * hpv
work_timing_inconvenient * hpv
too_far * hpv
no_reason * hpv
/ NOROW MISSING;
RUN;

/* Hesitancy only */
PROC FREQ DATA=survey_data_recode;
WHERE UPCASE(Vaccine_hesitancy) = "YES";
TABLES did_not_think_needed did_not_think_effective did_not_think_safe
not_enough_data bad_experience_vaccine bad_experience_clinic
provider_restriction other_beliefs religious_reasons
work_timing_inconvenient too_far_no_reason / NOROW MISSING;
RUN;

PROC FREQ DATA=survey_data_recode;
WHERE UPCASE(Vaccine_hesitancy) = "YES";
TABLES
did_not_think_needed * hpv
did_not_think_effective * hpv
did_not_think_safe * hpv
not_enough_data * hpv
bad_experience_vaccine * hpv
bad_experience_clinic * hpv
provider_restriction * hpv
other_beliefs * hpv
religious_reasons * hpv
work_timing_inconvenient * hpv
too_far * hpv
no_reason * hpv
/ NOROW MISSING;
RUN;

/* Refusal only */
PROC FREQ DATA=survey_data_recode;
WHERE UPCASE(Vaccine_refusal) = "YES";
TABLES did_not_think_needed did_not_think_effective did_not_think_safe
not_enough_data bad_experience_vaccine bad_experience_clinic
provider_restriction other_beliefs religious_reasons
work_timing_inconvenient too_far_no_reason / MISSING;
RUN;

PROC FREQ DATA=survey_data_recode;
WHERE UPCASE(Vaccine_refusal) = "YES";
TABLES
did_not_think_needed * hpv
did_not_think_effective * hpv
did_not_think_safe * hpv
not_enough_data * hpv
bad_experience_vaccine * hpv
bad_experience_clinic * hpv
provider_restriction * hpv
other_beliefs * hpv
religious_reasons * hpv
work_timing_inconvenient * hpv
too_far * hpv
no_reason * hpv
/ NOROW MISSING;
RUN;

/****** LOGISTIC REGRESSIONS - HPV vs attitudes/scores***** */
/* unadjusted & adjusted */
PROC LOGISTIC DATA=survey_data_recode DESCENDING;
CLASS Vaccine_Important (REF="Agree") / PARAM=REF;
MODEL hpv(event="No") = Vaccine_important;
RUN;

PROC LOGISTIC DATA=survey_data_recode DESCENDING;
CLASS Vaccine_Important (REF="Agree")
gender1 racel income education1 / PARAM=REF;
MODEL hpv(event="No") =
Vaccine_important age1 gender1 racel income education1;
RUN;

PROC LOGISTIC DATA=survey_data_recode DESCENDING;
CLASS Vaccine_Safe (REF="Agree") / PARAM=REF;
MODEL hpv(event="No") = Vaccine_Safe;
RUN;

PROC LOGISTIC DATA=survey_data_recode DESCENDING;
CLASS Vaccine_Safe (REF="Agree")
gender1 racel income education1 / PARAM=REF;
MODEL hpv(event="No") =
Vaccine_Safe age1 gender1 racel income education1;
RUN;

PROC LOGISTIC DATA=survey_data_recode DESCENDING;
CLASS Vaccine_Effective (REF="Agree") / PARAM=REF;
MODEL hpv(event="No") = Vaccine_Effective;
RUN;

PROC LOGISTIC DATA=survey_data_recode DESCENDING;
CLASS Vaccine_Effective (REF="Agree")
gender1 racel income education1 / PARAM=REF;
MODEL hpv(event="No") =
Vaccine_Effective age1 gender1 racel income education1;
RUN;

PROC LOGISTIC DATA=survey_data_recode DESCENDING;
CLASS Personal_Belief_Compatible (REF="Agree") / PARAM=REF;
MODEL hpv(event="No") =
Personal_Belief_Compatible age1 gender1 racel income education1;
RUN;

/****** Scores ***** */
/* Descriptive stats for scores */
PROC MEANS DATA=survey_data_recode N MEAN MEDIAN STD MIN Q1 Q3 MAX;
VAR score score2 scoreb;
RUN;

PROC MEANS DATA=survey_data_recode N MEAN MEDIAN STD MIN Q1 Q3 MAX MAXDEC=2;
CLASS hpv;
VAR score score2;
RUN;

PROC MEANS DATA=survey_data_recode MEDIAN MAXDEC=2;
VAR safe_score imp_score eff_score belief_score;
RUN;

/* score as continuous predictor */
PROC LOGISTIC DATA=survey_data_recode DESCENDING;
CLASS gender1 racel income education1 / PARAM=REF;
MODEL hpv(event="No") = score;
RUN;

PROC LOGISTIC DATA=survey_data_recode DESCENDING;
CLASS gender1 racel income education1 / PARAM=REF;
MODEL hpv(event="No") =
score age1 gender1 racel income education1;
RUN;

/* score2 as continuous predictor */
PROC LOGISTIC DATA=survey_data_recode DESCENDING;
CLASS gender1 racel income education1 / PARAM=REF;
MODEL hpv(event="No") = score2;
RUN;

PROC LOGISTIC DATA=survey_data_recode DESCENDING;
CLASS gender1 racel income education1 / PARAM=REF;
MODEL hpv(event="No") =
score2 age1 gender1 racel income education1;
RUN;

/* scoreb as binary predictor (cutoff=) */
PROC LOGISTIC DATA=survey_data_recode DESCENDING;
CLASS gender1 racel income education1 / PARAM=REF;
MODEL hpv(event="No") = scoreb;
RUN;

PROC LOGISTIC DATA=survey_data_recode DESCENDING;
CLASS gender1 racel income education1 / PARAM=REF;
MODEL hpv(event="No") =
scoreb age1 gender1 racel income education1;
RUN;

/* Distributions for attitude categories x scoreb */
PROC FREQ DATA=survey_data_recode;
TABLES Vaccine_Important Vaccine_Safe Vaccine_Effective
Personal_Belief_Compatible scoreb / MISSING;
RUN;

PROC FREQ DATA=survey_data_recode;
TABLES scoreb * hpv / MISSING NOROW CHISQ;
RUN;

/* Binary recode for vac questions */
DATA survey_data_recode2;
SET survey_data_recode;
safe_cat = (safe_score > 1); /* 0 = s1, 1 = s1 */
imp_cat = (imp_score > 1);
eff_cat = (eff_score > 1);
belief_cat = (belief_score > 1);
RUN;

PROC FORMAT;
VALUE bin_fmt
0 = "At or Below Median (<1)"
1 = "Above Median (>1)";
RUN;

PROC FREQ DATA=survey_data_recode2;
FORMAT safe_cat imp_cat eff_cat belief_cat bin_fmt.;
TABLES safe_cat imp_cat eff_cat belief_cat / NOCUM;
RUN;

PROC FREQ DATA=survey_data_recode2;
FORMAT safe_cat imp_cat eff_cat belief_cat bin_fmt.;
TABLES hpv * (safe_cat imp_cat eff_cat belief_cat)
/ NOCUM NOPERCENT CHISQ;
RUN;
```