Assignment 3

1. (a) $S \to U$, $V \to QU$, and $S \to TVX$ violate BCNF, since:
    - $S \to U$ is non-trivial and $S^+ = SUTVXQ$, meaning S is not a superkey and $S \to U$ violates BCNF.
    - $R \to SW$ is non-trivial and $R^+ = RSWUTVXQ$, meaning R is a superkey and $R \to SW$ does not violate BCNF.
    - $V \to QU$ is non-trivial and $V^+ = VQU$, meaning V is not a superkey and $V \to QU$ violates BCNF.
    - $S \to TVX$ is non-trivial and $S^+ = SUTVXQ$, meaning S is not a superkey and $S \to TVX$ violates BCNF.

   (b) I use $V \to QU$, which violates BCNF, to split R_1 into the relations $T_1(Q, U, V)$ and $T_2(R, S, T, V, W, X)$. To see if T_1 satisfies BCNF, I project the FDs onto it as follows:

Q	U	V	Closure	Projected FDs
✓			$Q^+ = Q$	Nothing
	✓		$U^+ = U$	Nothing
		✓	$V^+ = VQU$	$V \to QU$
✓	✓		$QU^+ = QU$	Nothing

I did not check the rest of the subsets of T_1 (which are QV, UV, and QUV) since they are all supersets of V, which is a superkey, and they can only generate weaker FDs than what I already have. Since T_1 does not have a projected FD that violates BCNF, it satisfies BCNF. To see if T_2 satisfies BCNF, I project the FDs onto it as follows:

R	S	T	V	W	X	Closure	Projected FDs
✓						$R^+ = RSWUTVXQ$	$R \to SWTVX$
	✓					$S^+ = SUTVXQ$	$S \to TVX$: violates BCNF; abort the projection

Since T_2 has a projected FD that violates BCNF, it does not satisfy BCNF and I must decompose it further. I use $S \to TVX$, the violating FD, to split T_2 into the relations $T_3(S, T, V, X)$ and $T_4(R, S, W)$. To see if T_3 satisfies BCNF, I project the FDs onto it as follows:

S	T	V	X	Closure	Projected FDs
✓				$S^+ = SUTVXQ$	$S \to TVX$
	✓			$T^+ = T$	Nothing
		✓		$V^+ = VQU$	Nothing
			✓	$X^+ = X$	Nothing
	✓	✓		$TV^+ = TVQU$	Nothing
	✓		✓	$TX^+ = TX$	Nothing
		✓	✓	$VX^+ = VXQU$	Nothing
	✓	✓	✓	$TVX^+ = TVXQU$	Nothing

I did not check the rest of the subsets of T_3 (which are STV, STX, SVX, and $STVX$) since they are all supersets of S, which is a superkey. Since T_3 does not have a projected FD that violates BCNF, it satisfies BCNF. To see if T_4 satisfies BCNF, I project the FDs onto it as follows:

R	S	W	Closure	Projected FDs
✓			$R^+ = RSW$	$R \to SW$
	✓		$S^+ = SUTVXQ$	Nothing
		✓	$W^+ = W$	Nothing
	✓	✓	$SW^+ = SWUTVXQ$	Nothing

I did not check the rest of the subsets of T_4 (which are RS, RW, and RSW) since they are all supersets of R, which is a superkey. Since T_4 does not have a projected FD that violates BCNF, it satisfies BCNF.

In summary, the relations in the final decomposition and their corresponding projected FDs are: $T_1(Q, U, V)$ with FD $V \to QU$, $T_3(S, T, V, X)$ with FD $S \to TVX$, and $T_4(R, S, W)$ with FD $R \to SW$.

(c) My schema preserves dependencies. Note that $S \to U$ follows from $S \to TVX$ and $V \to QU$. Although $S \to U$ is not a projected FD in the final decomposition, $S \to TVX$ and $V \to QU$ are projected, meaning $S \to U$ must also hold for any instance of the final decomposition. Hence, all dependencies are preserved in the schema.

(d) Assume $< q, r, s, t, u, v, w, x >$ came from $T_1 \bowtie T_3 \bowtie T_4$. Then, a possible instance of R_1 is:

Q	R	S	T	U	V	W	X
q	1	2	3	u	v	4	5
6	7	s	t	8	v	9	x
10	r	s	11	12	13	w	14

which violates the FDs $S \to U$, $V \to QU$, and $S \to TVX$. To make this instance of R_1 valid, I make the following corrections (the values to replace are marked with strikethroughs):

Q	R	S	T	U	V	W	X
q	1	2	3	u	v	4	5
~~6~~ q	7	s	t	~~8~~ u	v	9	x
~~10~~ q	r	s	~~11~~ t	~~12~~ u	~~13~~ v	w	~~14~~ x

Since $< q, r, s, t, u, v, w, x >$ appears in the last row of this instance, it must appear in R_1. Thus, the decomposition has lossless joins.

2. (a) I first split the RHS of each FD, and the set of resulting FDs is $S_3 = \{CDH \to F, G \to D, G \to H, FG \to C, FG \to D, FG \to E, H \to C, H \to E, H \to G, F \to C, F \to D\}$. Next, I try to reduce the LHS of FDs with multiple attributes on the LHS to make them stronger:

- $CDH \to F$: $C^+ = C$, $D^+ = D$, $H^+ = HCEGDF$, so I can reduce the LHS to H.
- $FG \to C$: $F^+ = FCD$, so I can reduce the LHS to F.
- $FG \to D$: $F^+ = FCD$, so I can reduce the LHS to F.
- $FG \to E$: $F^+ = FCD$, $G^+ = GDHCE$, so I can reduce the LHS to G.

The FDs after reducing their LHS are numbered as follows:

- 1: $H \to F$
- 2: $G \to D$
- 3: $G \to H$
- 4: $F \to C$
- 5: $F \to D$
- 6: $G \to E$
- 7: $H \to C$
- 8: $H \to E$
- 9: $H \to G$

I name this set of FDs S_4. Next, I find and eliminate redundant FDs:

FD	Exclude these from S_4 when computing closure	Closure	Decision
1	1	There's no way to get F without this FD	Keep
2	2	$H^+ = GHEFCD$	Discard
3	2, 3	There's no way to get H without this FD	Keep
4	2, 4	There's no way to get C without this FD	Keep
5	2, 5	There's no way to get D without this FD	Keep
6	2, 6	$G^+ = GHFCE$	Discard
7	2, 6, 7	$H^+ = HFEGCD$	Discard
8	2, 6, 7, 8	There's no way to get E without this FD	Keep
9	2, 6, 7, 9	There's no way to get G without this FD	Keep

No further simplifications are possible. The remaining FDs are 1, 3, 4, 5, 8, and 9, and the minimal basis is $S_5 = \{F \to C, F \to D, G \to H, H \to E, H \to F, H \to G\}$.

(b) I first determine which attributes must be in every key and which ones I must check:

Attribute	Appears on LHS	Appears on RHS	Conclusion
A, B	–	–	must be in every key
–	✓	–	must be in every key
C, D, E	–	✓	is not in any key
F, G, H	✓	✓	must check

Then, I check all possible combinations of F, G, and H added to A and B:

F	G	H	Closure	Decision
✓			$ABF^+ = ABFCD$	Not a superkey
	✓		$ABG^+ = ABGHEFCD$	Superkey
		✓	$ABH^+ = ABHEFGCD$	Superkey

I did not check the rest of the combinations (which are $ABFG$, $ABFH$, $ABGH$, and $ABFGH$) since they are all supersets of either ABG or ABH, which are superkeys. Also, notice that no subset of ABG and ABH are superkeys:

- $A^+ = A$, so A is not a superkey.
- $B^+ = B$, so B is not a superkey.
- $G^+ = GHEFCD$, so G is not a superkey.
- $H^+ = HEFGCD$, so H is not a superkey.
- $AB^+ = AB$, so AB is not a superkey.
- $AG^+ = AGHEFCD$, so AG is not a superkey.
- $BG^+ = BGHEFCD$, so BG is not a superkey.
- $AH^+ = AHEFGCD$, so AH is not a superkey.
- $BH^+ = BHEFGCD$, so BH is not a superkey.

Thus, ABG and ABH are minimal superkeys, meaning they are the keys for R_2.

(c) I first merge the RHS of the FDs in the minimal basis to form $S_6 = \{F \to CD, G \to H, H \to EFG\}$. Using this, I define the relations $R_3(F, C, D)$, $R_4(G, H)$, and $R_5(H, E, F, G)$. Since the attributes GH are in R_5, I don't need to keep R_4. Also, no relation is a superkey, so I add the relation $R_6(A, B, G)$, which includes the key ABG. The relations in the final decomposition are:

$$R_3(F, C, D) \qquad R_5(H, E, F, G) \qquad R_6(A, B, G).$$

(d) To check for redundancy, I must determine if there is any relation in the final decomposition that does not satisfy BCNF. To see if R_3 satisfies BCNF, I project the minimal basis onto it as follows:

F	C	D	Closure	Projected FDs
✓			$F^+ = FCD$	$F \rightarrow CD$
	✓		$C^+ = C$	Nothing
		✓	$D^+ = D$	Nothing
	✓	✓	$CD^+ = CD$	Nothing

I did not check the rest of the subsets of R_3 (which are FC, FD, and FCD) since they are all supersets of F, which is a superkey. Since R_3 does not have a projected FD that violates BCNF, it satisfies BCNF. To see if R_4 satisfies BCNF, I project the minimal basis onto it as follows:

H	E	F	G	Closure	Projected FDs
✓				$H^+ = HEFGCD$	$H \rightarrow EFG$
	✓			$E^+ = E$	Nothing
		✓		$F^+ = FCD$	Nothing
			✓	$G^+ = GHEFCD$	$G \rightarrow H$
	✓	✓		$EF^+ = EFCD$	Nothing

I did not check the rest of the subsets of R_4 (which are HE, HF, HG, EG, FG, HEF, HEG, HFG, EFG, and $HEFG$) since they are all supersets of either H or G, which are superkeys. Since R_4 does not have a projected FD that violates BCNF, it satisfies BCNF. To see if R_5 satisfies BCNF, I project the minimal basis onto it as follows:

A	B	G	Closure	Projected FDs
✓			$A^+ = A$	Nothing
	✓		$B^+ = B$	Nothing
		✓	$G^+ = GHEFCD$	Nothing
✓	✓		$AB^+ = AB$	Nothing
✓		✓	$AG^+ = AGHEFCD$	Nothing
	✓	✓	$BG^+ = BGHEFCD$	Nothing
✓	✓	✓	$ABG^+ = ABGHEFCD$	Nothing

Since R_5 does not have any projected FDs, it does not have a projected FD that violates BCNF, meaning it satisfies BCNF. In summary, R_3, R_5, and R_6 all satisfy BCNF. Therefore, none of them allows redundancy.