Business Analyst Interview Guide

Contents

Venkata Reddy Konasani
21.VENKAT@GMAIL.COM

Statistics -1

1. What is regression
   a. A mathematical equation between variables which reflects their dependency, can be used in prediction
2. What is correlation
   a. A coefficient which quantifies the linear relation
3. Difference between correlation and regression
   a. One is a model (equation) one is a coefficient
4. What is covariance
   a. Variance of one variable with respect to variance in other
5. What diff. between covariance and correlation
   a. Both are one and the same, corr is scaled version of cov (correlation is unit free, can be directly used in comparisons like corr(X1,Y1) > corr(X2,Y2), but it cann't be done directly with covariance)
6. Explain mean, mode and median
   a. Mean is average ie=sum/count (used when the data is concentrated)
   b. Mode is also average: Most occurring value (used for more discrete data OR categorical variables)
   c. Median is also average: Exact middle point of the data distribution (used as a replacement for mean when the data has outliers)
7. What is p-value
   a. It is the least significant level at which you can accept the null hypothesis.
   b. More simply, Probability of H0 is true; if it is higher H0 is accepted. When do u reject null hypothesis based on p-value
   c. If the observed statistic happens to be just a chance, p-values tells us what is the probability of that chance
   d. **The P-value answer the question:** What is the probability of the observed test statistic or one more extreme when H0 is true?
   e. Given H0, probability of the current value or extreme than this
   f. Given H_0 is true, probability of obtaining a result as extreme or more extreme than the actual sample
   g. The **observed significance level,** or **p-value** of a test of hypothesis is the probability of obtaining the observed value of the sample statistic, or one which is even more supportive of the alternative hypothesis, under the assumption that the null hypothesis is true.
   h. Smallest α the observed sample would reject H_0

8. What is MLE (maximum likelihood estimator)
   a. Used in estimating statistical parameters, It assumes a (NO) distribution of the parameter and maximizes its joint probability distribution, estimate is obtained at the point where probability distribution of parameter is maximum.
   b. Starts with arbitrary values of the regression coefficients and constructs an initial model for predicting the observed data. Then evaluates errors in such prediction and changes the regression coefficients so as make the likelihood of the observed data greater under the new model. Repeats until the model converges, meaning the differences between the newest model and the previous model are trivial.
   c. It assumes the distribution of the variable under consideration and finds out the parameters of that distribution by maximizing the likelihood function (JDF).
9. What is difference between likelihood function and joint distribution
   a. Joint distribution is function a function of parameters whereas likelihood function of sample observations
10. What is OLSE (ordinary least squares estimator)
    a. Minimizing the sum of squared deviation
11. Why do we minimize squares of deviations (OLSE) why cannot we use absolute differences
    a. It is hard to deal with absolute differences when you are differentiating and integrating.
12. What are the assumptions for using ordinary regression
    a. Errors normal, Errors iid
13. When do u go for generalized linear models (GLM)
    a. When above assumptions fail…….eg: errors are not normal;
    b. OR When you have discrete independent variable eg: yes /no; 1 or 0
14. How do u find parameters of GLM (OLSE or MLE)
    a. MLE
15. What is multi co-linearity, how do you detect it and how will you remove it?
    a. Interdependency of independent (predictor variables)
    b. VIF is used to detect, high F but low t values, conditional index
    c. Use PCA or FA, Drop problematic variables, Ridge regression
16. What is variance inflation factor (VIF)
    a. Each of the predictor variables is regressed up on other predictor vars. If that R-squared is high then this variable has collinearity with others.
17. Why do we check for multi co-linearity in nonlinear regression models?
    a. Model is nonlinear i.e the regression coefficients are nonlinear not predictor variables, so there is a chance of colinearity relation between the predictor vars.
18. What procedure do you use to fit regression model in SAS?
    a. Proc reg

19. What is the procedure for non-linear regression in SAS?
    a. Proc logistic, Proc genmod
20. How do u merge two data sets in SAS
    a. Merge statement
    b. Proc sql
21. What are the procedures for finding basic statistics (mean, median, skewness etc.) in SAS
    a. Proc univariate, Proc means
22. What is logistic regression, when do u use it?
    a. When basic assumption of regression fail,
    b. When there is binary(categorical) response
23. Give an example where you can use logistic regression
    a. Response / Non response
    b. Good customer, Bad customer……all binary cases
24. How do you measure goodness of your regression model?
    a. R-square.
25. What is log-linear model?
    a. Use log a link function instead of logit, which is used for poisson response variable (y = 0,1,2,3,…)
26. When do u use paired t-test?
    a. To test the difference between two dependent samples (basically one sample)
    b. Eg: before treatment and after treatment of a particular disease.
27. How do u find goodness of your model in GLM?
    a. Its not R-square, here it is Chi-square.
    b. Percent Correct Predictions
    c. Hosmer and Lemeshow Goodness-of-Fit Test
    d. ROC curves
    e. Somers' D
    f. Gamma
    g. Tau-a
    h. C
    i. More than a dozen "R2"-type summaries
28. What criteria do u use to find parameters of ordinary regression?
    a. least squares method
29. When do you use t-test and what are assumptions there?
    a. To test the small sample mean….whether it is equal to population mean or not

SAS-1

1. What are some of the differences between a WHERE and an IF statement?
   - IF can only be used in a DATA step
   - Many IF statements can be used in one DATA step
   - Must read the record into the program data vector to perform selection with IF
   - WHERE can be used in a DATA step as well as a PROC.
   - A second WHERE will replace the first unless the option ALSO is used
   - Data subset prior to reading record into PDV with WHERE

2. Name some of the ways to create a macro variable
   a. %let
   b. CALL SYMPUT (....)
   c. when creating a macro example:
   d. %macro mymacro(a=,b=);
   e. in SQL using INTO

3. What is the difference between proc means and proc univariate?
   a. Both procedure produce descriptive statistics. By proc univariate, by default it produces all the statistics (some time not all are required) but in proc means it is possible to request the statistics that we want.
   b. With proc univariate you can generate histograms (graphs) for particular variables. You can't do that with proc means.

4. What is the difference between an informat and a format? Name three informats or formats.
   a. Informat to read the data. Format to write the data.

5. How would you delete observations with duplicate keys?
   a. use PROC SORT with the option NODUPKEY

6. Which program displays a listing of all data sets in the SASUSER library?
   a. proc contents data = sasuser._all_; run;

7. What is the maximum character of SAS library name?
   a. A valid library name must start with an alphabet and cannot have more than **8** characters.

8. What are the most commonly specified options in SAS system?
   a. errors = n; compress = yes, reuse = yes, obs = n, nodate, mlogic, mprint, symbolgen

9. What are the common SAS log notes and warnings?
   a. NOTE: Missing values were generated as a result of performing an operation on missing values.
   b. NOTE: MERGE statement has more than one data set with repeats of BY values.
   c. WARNING: Multiple lengths were specified for the BY variable emp_id by input data sets. This may cause unexpected results.
   d. WARNING: The variable x in the DROP, KEEP, or RENAME list has never been referenced
   e. NOTE: Variable var1 is uninitialized.

f. NOTE: Numeric values have been converted to character values at the places given by: (Line):(Column). 52:8

g. NOTE: Character values have been converted to numeric values at the places given by: (Line):(Column).55:9

10. What is the function to convert variables from numeric to character and character to numeric?

   a. Convert from numeric to character with the PUT function Character-variable = put (numeric-variable, format.);

   b. Convert from character to numeric with the INPUT function Numeric-variable = input (character-variable, informat.);

11. How is Proc SQL different from SAS?

   a. The SQL procedure is a very effective method of sorting, counting, and selecting data that does not require the same amount of processing resources or time as traditional procedures

   b. SQL looks at datasets differently from SAS

      i. SAS looks at a dataset one record at a time, using an implied loop that moves from the first record to the last

      ii. SQL looks at all the records, as a single object

      iii. Because of this difference SQL can easily do a few things that are more difficult to do in SAS PROC SQL is an additional tool with its own strengths and challenges

   c. The SQL procedure can be a very efficient tool to use when processing very large data sets. It also minimizes the number of steps that need to be written to perform a task. SQL saves processing time, programmer time and memory used.

12. SQL - Combining Datasets: Joins

   a. Full join, Inner join, Left join, Right join, and Cartesian join

   b. Proc SQL; Create table d3 as Select * From d1 Full join d2 On d1.var=d2.var; Quit;

   c. SQL allows merging (joining) where key variables have different names

   d. SQL does not require sorting

13. Can you tell us the commonly base procedure that you used in last one year?

   a. **Report** Writing Procedures - PRINT, FREQ, MEANS, SUMMARY, TABULATE, PLOT, SQL

   b. **Statistical** Procedures – CHART, FREQ, MEANS, CORR, SQL, SUMMARY, UNIVARIATE

   c. **Utility** Procedures **–** EXPORT, IMPORT, APPEND, CONTENTS, DATASETS, SORT, TRANSPOSE

14. Can you give some example of character functions like Substr and Scan?

   a. Syntax: X = substr(argument,position,n)

   b. Y = "substring" , "developer"   X = substr(Y,4,3) , substr(Y,4,5)

   c. X = "str" , "elope"

   d. Syntax: X = scan(argument,n,delimiters)

   e. Y = "C and Y" , "dilligent"   X = scan(Y,2," ") , scan(Y,3,"i")

   f. X = "and" , "gent"

15. What happens if you don't use a 'BY' statement in your MERGE?
    a. While merging if you don't use a 'BY' statement; SAS will merge the 1st observation from dataset A and 1st observation from dataset B to form the 1st observation of the final dataset
    b. Do not miss out on the 'BY' statement. It might lead to useless results.
16. List differences between SAS PROCs and the SAS DATA STEP.
    a. Procs are sub-routines with a specific purpose in mind and the data step is designed to read in and manipulate data
17. What is the order of select statement in SQL?
    a. When you construct a SELECT statement, you must specify the clauses in the following order:
       o SELECT
       o FROM
       o WHERE
       o GROUP BY
       o HAVING
       o ORDER BY
18. How do you eliminate duplicate rows from the query results?
    a. You can eliminate the duplicate rows from the results by using the DISTINCT keyword in the SELECT clause.
19. How to drop or keep columns in SQL?
    a. Through SELECT
20. How to rename column in SQL?
    a. Through AS
21. What is the use of CASE expression in SQL?
    a. You can use conditional logic within a query by using a CASE expression to conditionally assign a value. You can use a CASE expression anywhere that you can use a column name.

Statistics -2

1. What do you mean by Central tendency?

If all the homogeneous data plotted in graphically it is observed that data tends to cluster in one region which will have a midpoint which can denote the overall data. This characteristic is called as central tendency.

1.2. What is variance?

It is root mean square distance of all the data from central point (Mean, Median, Mode).

1.3. Difference between Mean, Median, and Mode based on the type of average?

Mean is arithmetic average, Median is positional average, Mode is frequencial average.

2. What is sample and population?

Sample is a subset of the population which represents the population. It will have all characteristics of the population. All possible observations under study is called population.

 Specific characteristics of a sample?

- It should be subset of population.
- It will have all characteristics of the population

Different types of probability sampling

- Simple random sampling
- Stratified sampling
- Cluster sampling
- Systematic sampling

3. What is a cluster in sampling?

While doing cluster sampling we divide the population into groups which are homogenous within the group and heterogeneous between the groups. These units are called cluster.

The technique works best when most of the variation in the population is within the groups, not between them.

4. What is the difference between stratified and cluster sampling?

Basically in a stratified sampling procedure, the population is first partitioned into disjoint classes (the strata) which together are exhaustive. Thus each population element should be within one and only one stratum. Then a simple random sample is taken from each stratum, the sampling effort may either be a proportional allocation (each simple random sample would contain an amount of variates from a stratum which is proportional to the size of that stratum) or according to optimal allocation, where the target is to have a final sample with the minimum variabilty possible.

The main difference between stratified and cluster sampling is that in stratified sampling all the strata need to be sampled. In cluster sampling one proceeds by first selecting a number of clusters at random and then sampling each cluster or conduct a census of each cluster. But usually not all clusters would be included.

5. How do you determine the size of sample?

- With experience
- Depending on requirement of the project
- Client request
- Time and cost budget
- Statistical formula for minimum sample size

6. Type I and Type II error?

In a statistical hypothesis test, there are two types of incorrect conclusions that can be drawn. The hypothesis can be inappropriately rejected (this is called type I error), or one can inappropriately retain the hypothesis (this is called type II error). The Greek letter α is used to denote the probability of type I error, and the letter β is used to denote the probability of type II error.

Rejecting a null-hypothesis when it should not have been rejected creates a type I error - failing to reject a null-hypothesis when it should have been rejected creates a type II error.

Power of the test is 1- β

7. What is Normal distribution?

It is a continuous symmetric distribution for which Mean= Median= Mode. It is symmetric distribution that's why it is normal. Any distribution tends to normal if no of observation tends to infinity.

8. What is Poisson distribution?

In probability theory and statistics, the Poisson distribution (or Poisson law of small numbers) is a discrete probability distribution that expresses the probability of a number of events occurring

in a fixed period of time if these events occur with a known average rate and independently of the time since the last event. (The Poisson distribution can also be used for the number of events in other specified intervals such as distance, area or volume.) Chances of the fevourable out come should be high (rare event).

9. What is Central Limit theorem?

In probability theory , the central limit theorem (CLT) states conditions under which the mean of a sufficiently large number of independent random variables , each with finite mean and variance, will be approximately normally distributed (Rice 1995). The central limit theorem also requires the random variables to be identically distributed, unless certain conditions are met. Since real-world quantities are often the balanced sum of many unobserved random events, this theorem provides a partial explanation for the prevalence of the normal probability distribution. The CLT also justifies the approximation of large-sample statistics to the normal distribution in controlled experiments.

10. What is correlation?

It is a measure of degree of association between two related variable. X is directly proportional to Y is called positive correlation. X is inversely proportional to Y is called negative correlation.

   a. If X takes value 1 2 3 4 5 what could be the value of Y if the corelation coefficient of X&Y = 0?

   Y will be constant or Zero.

   b. If correlation coefficient is 1.1 between demand and supply what would you understand from this?

   The calculation is wrong. $-1 <= R <= 1$

11. When can you do T-test instead of Z-test?

When sample size is less than 30, we do t-test.

12. When can you do parametric test and when can you do non parametric test based on population?

When the population is defined we go for parametric test as we can calculate probability, mean, stdv of the population and when the population is not defined we go for non parametric test.

13. What is regression analysis?

Regression analysis is a statistical technique for estimating the relationships among variables. It basically try to measure and identify the cause and effect relationship among the variables. Regression analysis comes in different flavors: Logistic, Multiple choice logistic, multinomial, multiple etc.

14. What is Multicollinearity?

Multicollinearity refers to a situation in which two or more explanatory variables in a multiple regression model are highly correlated. We have perfect multicollinearity if the correlation between two independent variables is equal to 1 or -1. In practice, we rarely face perfect multicollinearity in a data set. More commonly, the issue of multicollinearity arises when there is a high degree of correlation (either positive or negative) between two or more independent variables.

Multicollinearity is a statistical phenomenon in which two or more predictor variables in a multiple regression model are highly correlated. In this situation the coefficient estimates may change erratically in response to small changes in the model or the data. Multicollinearity does not reduce the predictive power or reliability of the model as a whole; it only affects calculations regarding individual predictors. That is, a multiple regression model with correlated predictors can indicate how well the entire bundle of predictors predicts the outcome variable, but it may not give valid results about any individual predictor, or about which predictors are redundant with others.

15. What is the difference b/w R2 and Adjusted R2?

R2 is a statistic that will give some information about the goodness of fit of a model. In regression, the R2 coefficient of determination is a statistical measure of how well the regression line approximates the real data points. An R2 of 1.0 indicates that the regression line perfectly fits the data.

R^2 increases whenever we add a new independent variable, Adjusted R^2 might increase or decrees based on the variable explanatory power. Too many independent variables can decrease the value of Adj R^2

Adjusted R2 is a modification of R2 that adjusts for the number of explanatory terms in a model. Unlike R2, the adjusted R2 increases only if the new term improves the model more than would be expected by chance. The adjusted R2 can be negative, and will always be less than or equal to R2. Adjusted R2 is not always better than R2: adjusted R2 will be more useful only if the R2 is calculated based on a sample, not the entire population. For example, if our unit of analysis is a state, and we have data for all counties, then adjusted R2 will not yield any more useful information than R2.

16. What is Chi-square test?

Chi-square is a statistical test commonly used to compare observed data with data we would expect to obtain according to a specific hypothesis. Chi-square tests are mainly of two types:

i) Goodness of fit

ii) Independence of attributes

17. What is ANOVA?

Analysis of variance between two or more variables. There are different types of anova like One-way, two-way, multiple. The ANOVA is based on the fact that two independent estimates of the population variance can be obtained from the sample data. A ratio is formed for the two estimates, where:

one is sensitive to ®	<u>treatment effect & error</u>	<u>between groups estimate</u>
and the other to ®	error	within groups estimate

Given the null hypothesis (in this case H_O: $m_1 = m_2 = m_3$), the two variance estimates should be equal. That is, since the null assumes no treatment effect, both variance estimates reflect error and their ratio will equal 1. To the extent that this ratio is larger than 1, it suggests a *treatment effect* (i.e., differences between the groups).

18. What is a hypothesis?

A hypothesis is a proposed explanation for an observable phenomenon. For a hypothesis to be put forward as a scientific hypothesis, the scientific method requires that one can test it. Scientists generally base scientific hypotheses on previous observations that cannot satisfactorily be explained with the available scientific theories. Even though the words "hypothesis" and "theory " are often used synonymously in common and informal usage, a scientific hypothesis is not the same as a scientific theory . A working hypothesis is a provisionally accepted hypothesis.

In a related but distinguishable usage, the term hypothesis is used for the antecedent of a proposition; thus in proposition "If P, then Q ", P denotes the hypothesis (or antecedent); Q can be called a consequent. P is the assumption in a (possibly counterfactual) What If question.

The adjective hypothetical, meaning "having the nature of a hypothesis," or "being assumed to exist as an immediate consequence of a hypothesis," can refer to any of these meanings of the term "hypothesis."

19. What is level of confidence (i.e. what 95% level significance stand for)?

Statistics "significant" means probably true (not due to chance). A research finding may be true without being important. When statisticians say a result is "highly significant" they mean it is very probably true. They do not (necessarily) mean it is highly important.

The amount of evidence required to accept that an event is unlikely to have arisen by chance is known as the significance level or critical p-value: in traditional Fisherian statistical hypothesis testing, the p-value is the probability conditional on the null hypothesis of the observed data or more extreme data. If the obtained p-value is small then it can be said either the null hypothesis is false or an unusual event has occurred. It is worth stressing that p-values do not have any repeat sampling interpretation.

20. Different kind of multivariate analysis?

- Factor analysis.
- Cluster analysis
- Conjoint analysis
- Discremenant analysis
- Logistic regression
- Multiple regressions

21. What is factor analysis?

Factor analysis is used to uncover the latent structure (dimensions) of a set of variables. It reduces attribute space from a larger number of variables to a smaller number of factors and as such is a "non-dependent" procedure (that is, it does not assume a dependent variable is specified). Factor analysis could be used for any of the following purposes:

- To reduce a large number of variables to a smaller number of factors for modeling purposes, where the large number of variables precludes modeling all the measures individually. As such, factor analysis is integrated in structural equation modeling (SEM), helping confirm the latent variables modeled by SEM. However, factor analysis can be and is often used on a stand-alone basis for similar purposes.
- To establish that multiple tests measure the same factor, thereby giving justification for administering fewer tests. Factor analysis originated a century ago with Charles Spearman's attempts to show that a wide variety of mental tests could be explained by a single underlying intelligence factor (a notion now rejected, by the way).

- To validate a scale or index by demonstrating that its constituent items load on the same factor, and to drop proposed scale items which cross-load on more than one factor.
- To select a subset of variables from a larger set, based on which original variables have the highest correlations with the principal component factors.
- To create a set of factors to be treated as uncorrelated variables as one approach to handling multicollinearity in such procedures as multiple regression
- To identify clusters of cases and/or outliers.
- To determine network groups by determining which sets of people cluster together (using Q-mode factor analysis, discussed below)

22. What are the different types of rotation in Factor loading?

Varimax rotation is an orthogonal rotation of the factor axes to maximize the variance of the squared loadings of a factor (column) on all the variables (rows) in a factor matrix, which has the effect of differentiating the original variables by extracted factor. Each factor will tend to have either large or small loadings of any particular variable. A varimax solution yields results which make it as easy as possible to identify each variable with a single factor. This is the most common rotation option.

Quartimax rotation is an orthogonal alternative which minimizes the number of factors needed to explain each variable. This type of rotation often generates a general factor on which most variables are loaded to a high or medium degree. Such a factor structure is usually not helpful to the research purpose.

Equimax rotation is a compromise between Varimax and Quartimax criteria.

Direct oblimin rotation is the standard method when one wishes a non-orthogonal (oblique) solution – that is, one in which the factors are allowed to be correlated. This will result in higher eigenvalues but diminished interpretability of the factors. See below.

Promax rotation is an alternative non-orthogonal (oblique) rotation method which is computationally faster than the direct oblimin method and therefore is sometimes used for very large datasets.

23. Tell us something about cluster analysis?

Cluster analysis or clustering is the assignment of a set of observations into subsets (called clusters) so that observations in the same cluster are similar in some sense. Clustering is a method

of unsupervised learning, and a common technique for statistical data analysis used in many fields, including machine learning, data mining, pattern recognition, image analysis and bioinformatics.

24. What is discreminant analysis?

The main purpose of a discriminant function analysis is to predict group membership based on a linear combination of the interval variables. The procedure begins with a set of observations where both group membership and the values of the interval variables are known. The end result of the procedure is a model that allows prediction of group membership when only the interval variables are known. A second purpose of discriminant function analysis is an understanding of the data set, as a careful examination of the prediction model that results from the procedure can give insight into the relationship between group membership and the variables used to predict group membership.

Market Research

1. What is the process of market research?

Problem definition -> Negotiation on budget with client -> Qualitative research (interview brain storming)-> Identifying parameters and building Questionnaire if primary research/ identifying source of data if secondary research-> Identifying population and design of sample-> Preparing analysis plan-> data collection->coding->verification & validation->Analysis of data->Interpretation->Presentation

2. Do you come across Open ended; close ended; semi open ended coding?

Assigning a different numerical value to parameters of estimate is coding.

No statistical tool can analyze data without numbers. (String /character/statement:-> numerical value.)

3. What are the skip patterns in questionnaire?

If one question is yes then skip some question and answer some set of question. Like if you are using brand A you will answer some set of question if Brand B then some other set of question.

4. What all major analysis performed in Retail market research?

Brand shifting, brand tracking, brand positioning, price analysis, marketing mix modeling, customer segmentation, opportunity for new product identification, Concept & product test.

5. What is Market Basket Analysis?

The term market basket analysis in the retail business refers to research that provides the retailer with information to understand the purchase behavior of a buyer. This information will enable the retailer to understand the buyer's needs and rewrite the store's layout accordingly, develop cross-promotional programs, or even capture new buyers (much like the cross-selling concept). An early illustrative example for this was when one super market chain discovered in its analysis that customers that bought diapers often bought beer as well, have put the diapers close to beer coolers, and their sales increased dramatically.

6. I want to know for my product what my target population is. How do I do that using which statistical technique?

Need to do market segmentation using Cluster analysis. First do hierarchical clustering to identify no of segment of the market for similar product and then do K-means cluster to identify the levels of the parameters of the cluster to target.

7. Why do we do Factor analysis?

To reduce the number of variables for a better presentation of the key factors.

8. Usually how the analysis happens in a customer engagement analysis etc whether we do factor analysis first or cluster analysis or regression?

First we do factor analysis to reduce variables then with the factors we do regression to identify important factors then we do Cluster analysis based on the important factors.

9. Why do we need weighting in market research?

To project characteristic with a sample for the population we have many constrains such as availability of proper sample. Like in population male female ration is 50:50. but the sample shoes 80:20 then we need to put a weight to make it in same proportion.

10. What are the benefits of online research?

Online research offers greater access and a substantial time and cost savings over phone, in-person, and mail data collection methods. You can also more quickly receive results with online research.

12. Will the quality of graphics accurately represent product concept, product, or advertising?

With the assistance of several major clients who extensively use graphics in their online research, SPSS Inc. has spent the past two years developing "best-in-class" graphic capabilities. We continue to invest in R&D in this area to ensure we maintain our high-quality standards. A caveat: graphics quality largely depends on the original image and the end-user's equipment.

17. My survey is very long. Can respondents take a break and then return to complete the survey at a later time?

Yes. Respondents can log off of a survey, log in again, and return to the survey at the exact page where they left off. This enables respondents to finish the survey at their leisure and helps improve completion rates.

18. What kind of experience does your staff have with survey research and online surveys?

Our staff has numerous years of professional research experience from both the client and supplier perspectives. Staff members have experience programming a wide range of online survey projects, including:

- Concept/brand testing
- Conjoint studies
- Profiling and segmentation
- Web site evaluations
- Customer satisfaction
- Brand image, awareness, and usage
- Commercial/TV program testing
- Print ad testing
- Multimedia evaluations
- Tracking studies
- Forecasting
- Business-to-consumer/
- business-to-business studies

19. Can you program any advanced statistical techniques that require a dynamic/fluid survey experience?

Yes. We currently offer discrete choice modeling, a type of conjoint analysis.

Excel

1. Concatenation

2. If condition (if, then, else): 7 nested ifs possible

IF(CONDITION, THEN_DO_THIS, ELSE_DO_THIS)

3. VLOOKUP &HLOOKUP

vlookup (lookup value, table name or range, column index, true/false)

True is given only when data is in ascending order. Otherwise give false (for text, always give false)

If vlookup can't find lookup value, and range lookup is TRUE, it uses the largest value that is less than or equal to lookup value. The leftmost column should have unique values, otherwise only the first row is captured

Can have nested vlookups

To give name to a range of cells (this name is unique in the whole workbook): select and in left box at the top, enter the name; to access the list (for deletion later on), press control + F3

HLOOKUP - (lookup value, table name or range, row index, true/false)

4. Conditional formatting: Format -> conditional formatting (for ex. Changing colours based on particular condition)

5. Data Validation: 3 types of error messages depending on what user does – stop, warning and information; this prevents user from entering junk values. Data -> Validation (note: if give a list, then a drop-down box comes from which the user can choose)

6. Sumif (range, criteria, cells_to_sum)

Countif (range, criteria)

7. Pivot Tables:

Consolidates data in the form of a table; go to Data -> Pivot table and pivot chart report; if there are multiple tables in the worksheet, go for Multiple consolidation ranges, otherwise select

Microsoft office excel list or database; select where you want to put the pivot table, and then drag the fields to the respective places for row, column and the data item you want to consolidate

After consolidation, data can be grouped (in both rows and columns) and you can provide names for those groups; you can double-click on the page fields in the top row and display the sum for the groups as well

If source data changes, you can right-click inside the pivot table and Refresh the data; the cells inside the pivot table can also be referred to, in other cells so that everything is linked together

Vlookup and hlookup are very useful in a pivot table (can select the table and name the table on the top left box for referencing by table name later on); if row contains countries, and column contains products and you have to access the number of details for a particular product in a particular country, form another table with country name and its column number and use nested lookup to access the details for a particular country's particular product

SAS-2

1) What is the effect of the OPTIONS statement ERRORS=1?
2) What's the difference between **VAR A1 - A4** and **VAR A1 — A4**?
3) What do the SAS log messages *"numeric values have been converted to character"* mean? What are the implications?
4) Why is a STOP statement needed for the POINT= option on a SET statement?
5) How do you control the number of observations and/or variables read or written?
6) Approximately what date is represented by the SAS date value of 730?
7) How would you remove a format that has been permanently associated with a variable??
8) What does the RUN statement do?
9) Why is SAS considered self-documenting?
10) What areas of SAS are you most interested in?
11) Briefly describe 5 ways to do a "table lookup" in SAS.
12) What versions of SAS have you used (on which platforms)?
13) What are some good SAS programming practices for processing very large data sets?
14) What are some problems you might encounter in processing missing values? In Data steps? Arithmetic? Comparisons? Functions? Classifying data?
15) How would you create a data set with 1 observation and 30 variables from a data set with 30 observations and 1 variable?
16) What is the different between functions and PROCs that calculate the same simple descriptive statistics?
17) If you were told to create many records from one record, show how you would do this using arrays and with PROC TRANSPOSE?
18) What are _numeric_ and _character_ and what do they do?
19) How would you create multiple observations from a single observation?
20) For what purpose would you use the RETAIN statement?
21) What is a method for assigning first.VAR and last.VAR to the BY group variable on unsorted data?
22) What is the order of application for output data set options, input data set options and SAS statements?
23) What is the order of evaluation of the comparison operators: + - * / ** ( ) ?
24) How could you generate test data with no input data?
25) How do you debug and test your SAS programs?
26) What can you learn from the SAS log when debugging?
27) What is the purpose of _error_?
28) How can you put a "trace" in your program?
29) Are you sensitive to code walk-throughs, peer review, or QC review?

30) Have you ever used the SAS Debugger?

31) What other SAS features do you use for error trapping and data validation?

32) How does SAS handle missing values in: assignment statements, functions, a merge, an update, sort order, formats, PROCs?

33) How many missing values are available? When might you use them?

34) How do you test for missing values?

35) How are numeric and character missing values represented internally?

36) What SAS statements would you code to read an external raw data file to a DATA step?1.   infile statement

37) How do you read in the variables that you need?

38) Are you familiar with special input delimiters? How are they used? dlm option in infile statement, typically a comma, or hyphen or space …

39) If reading a variable length file with fixed input, how would you prevent SAS from reading the next record if the last variable didn't have a value? use colan

40) What is the difference between an informat and a format? Name three informats or formats.

41) Name and describe three SAS functions that you have used, if any?

42) How would you code the criteria to restrict the output to be produced?

43) What is the purpose of the trailing @? The @@? How would you use them?

44) Under what circumstances would you code a SELECT construct instead of IF statements?

45) What statement do you code to tell SAS that it is to write to an external file? What statement do you code to write the record to the file?

46) If reading an external file to produce an external file, what is the shortcut to write that record without coding every single variable on the record?

47) If you're not wanting any SAS output from a data step, how would you code the data statement to prevent SAS from producing a set?

48) What is the one statement to set the criteria of data that can be coded in any step?

49) Have you ever linked SAS code? If so, describe the link and any required statements used to either process the code or the step itself.

50) How would you include common or reuse code to be processed along with your statements?

51) When looking for data contained in a character string of 150 bytes, which function is the best to locate that data: scan, index, or indexc?

52) If you have a data set that contains 100 variables, but you need only five of those, what is the code to force SAS to use only those variable?

53) Code a PROC SORT on a data set containing State, District and County as the primary variables, along with several numeric variables.

54) How would you delete duplicate observations?

55) How would you delete observations with duplicate keys?

56) How would you code a merge that will keep only the observations that have matches from both sets.
57) How would you code a merge that will write the matches of both to one data set, the non-matches from the left-most data set to a second data set, and the non-matches of the right-most data set to a third data set.
58) What is the Program Data Vector (PDV)? What are its functions?
59) Does SAS 'Translate' (compile) or does it 'Interpret'? Explain.
60) At compile time when a SAS data set is read, what items are created?
61) Name statements that are recognized at compile time only?
62) Identify statements whose placement in the DATA step is critical.
63) Name statements that function at both compile and execution time.
64) Name statements that are execution only.
65) In the flow of DATA step processing, what is the first action in a typical DATA Step?
66) What is _n_?

Data Audit & Data Sanitization

Data validation using basic contents

- Total number of observations
- Total number of fields
- Each field name, Field type, Length of field
- Format of field, Label

Basic Contents –Check points

- Are all variables as expected (variables names)
- Are there some variables which are unexpected say q9 r10?
- Are the data types and length across variables correct
- For known variables is the data type as expected (For example if age is in date format something is suspicious)
- Have labels been provided and are sensible
- If anything suspicious we can further investigate it and correct accordingly

SAS Code :

```
proc contents data=<<data name>>; run;
Useful options :
```

Short – Outputs the list of variables in a row by row format.
        Code : proc contents data=test **short**; run;

Out=_filename_ - Creates a data set wherein each observation is a variable

Data validation- Snapshot

- Printing the first few observations   all fields in the data set .It helps in better understanding of the variable by looking at it's assigned values.

Checkpoints for data snapshot output:

1. Do we have any unique identifier? Is the unique identifier getting repeated in different records?

2. Do the text variables have meaningful data?(If text variables have absurd data as '&^%*HF' then either the variable is meaningless or the variable has become corrupt or wasn't properly created.)

3. Are there some coded values in the data?(if for a known variable say State we have category codes like 1-52 then we need definition of how they are coded.)

4. Do all the variables appear to have data? (In case variables are not populated with non missing meaningful value it would show in print. We can further investigates using means statistics.)

SAS code :

proc print data=<<data set>>; run;

Useful options :

proc print data=<<data set>> label noobs heading=vertical;

var <<variable-list>>; by var1; run;

- Label:The label option uses variable labels as column headings rather than variable names (the default).

- Obs : Restricts the number of observations in the output

- Nobs: It omits the OBS column of output.

- Heading=vertical: It prints the column headings vertically. *This is useful when the names are long but the values of the variable are short.*

- Var: Specifies the variables to be listed and the order in which they will appear.

- By: By statement produces output grouped by values of the mentioned variables

Categorical field frequencies

- Calculate frequency counts cross-tabulation frequencies for Especially for categorical, discrete & class fields

- Frequencies

    - help us understanding the variable by looking at the values it's taking and data count at each value.

    - They also helps us in analyzing the relationships between variables by looking at the cross tab frequencies or by looking at association

Checkpoints for looking frequency table

1. Are values as expected?

2. Variable understanding : Distinct values of a particular variable, missing percentages

3. Are there any extreme values or outliers?

4. Any possibility of creating a new variable having small number of distinct category by clubbing certain categories with others.

SAS code:

Proc FREQ data =<dataset > <options> ;

TABLES requests < / options > ; // Gives Frequency Count or Cross Tab

BY <varl> ;   // Grouping output based on varl

WEIGHT variable < / option > ; //Specifying Weight (if applicable)

OUTPUT < OUT=SAS-data-set > options ; //Output results to another data set   run;

- Useful options :

  - Order=Freq - sorts by descending frequency count (default is the unformatted value). Ex: proc freq data=test order=freq; tables X1-X5; run;

  - Nocol/Norow/Nopercent - suppresses printing of column, row and cell percentages respectively of a cross tab. Ex : proc freq data=test; tables AGE*bad/nocol norow nopercent missing; run;

  - Missing- interprets missing values as non-missing and includes them in % and statistics calculations   ex : proc freq data=test; tables CHANNEL* BAD /missing; run;

  - Chisq - performs several chi-square tests. Ex: proc freq data=test; tables channel*bad/chisq; run;

Descriptive Statistics for continuous fields

- Distribution of numeric variables by calculating

  - N – Count of non missing observations

  - Nmiss – Count of Missing observations

  - Min, Max, Median, Mean

  - Quartile numbers & percentiles– P1, p5,p10,q1(p25),q3(p75), p90,p99

  - Stddev

  - Var

  - Skewness

- Kurtosis

Check List

- Are variable distribution as expected.
- What is the central tendency of the variable? Mean, Median and Mode across each variable
- Is the concentration of variables as expected ? What are quartiles?
- Indicates variables which are unary I.e stddev=0 ; the variables which are useless for the current objective.
- Are there any outliers / extreme values for the variable?
- Are outlier values as expected or they have abnormally high values -for ex for Age if max and p99 values are 10000. Then should investigate if it's the default value or there is some error in data
- What is the % of missing value associated with the variable?

SAS Code :

PROC UNIVARIATE data=<dataset>;

VAR *variable(s)*; run;

- Useful options :

  - PROC UNIVARIATE data=<dataset> plot normal;
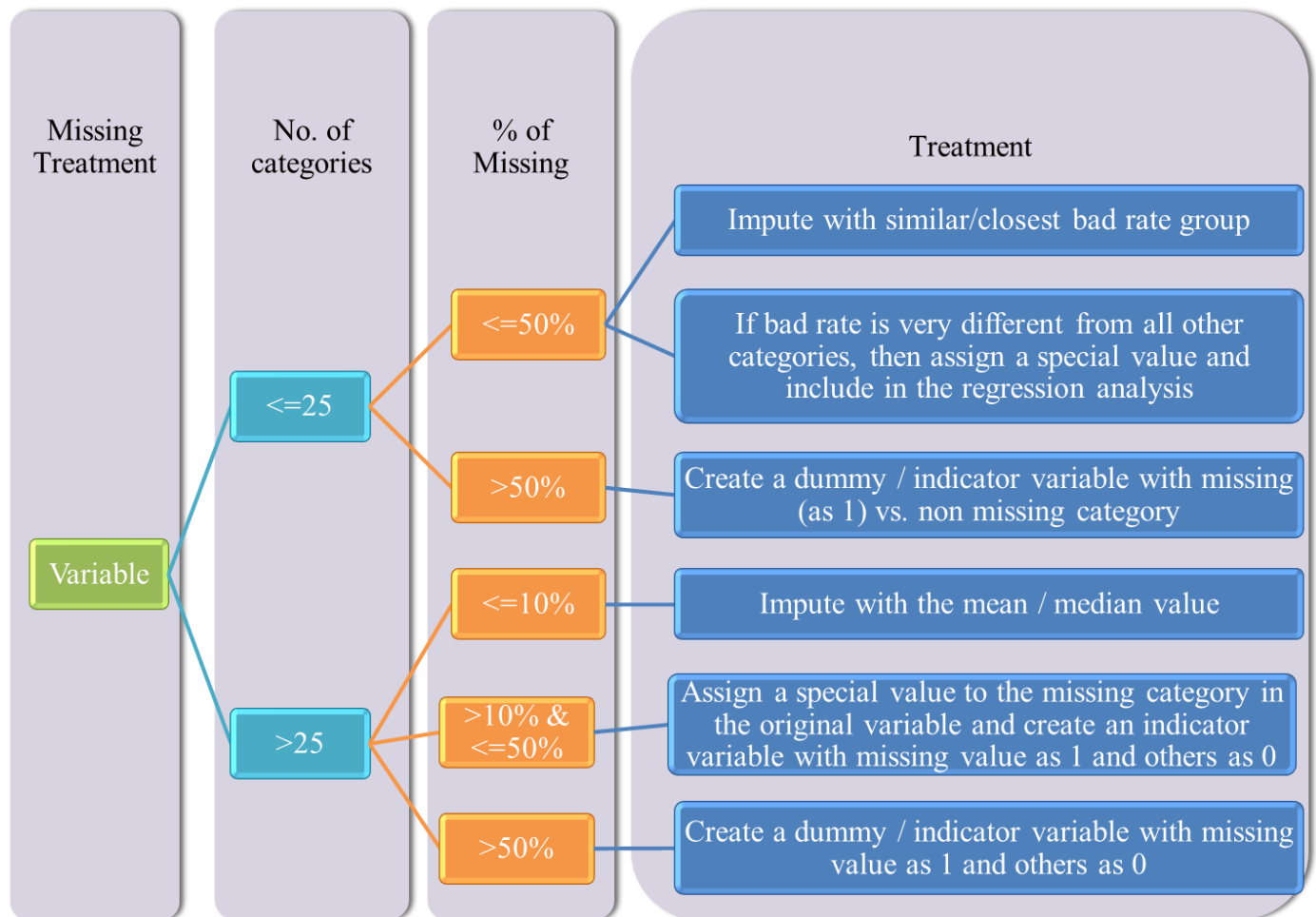
  HISTOGRAM *<variable(s)> </ option(s)>*;

  By variable;

  VAR *variable(s)*; run;

  - Normal option produces the tests of normality ;

  - Plot option produces the 3 plots of data(stem and leaf plot, box plot, normal probability plot

  - By option is used for giving outputs separated by categories

  - Histogram option gives the distribution of variable in a histogram

Missing Value & outlier Treatment

Missing Treatment	No. of categories	% of Missing	Treatment
Variable	<=25	<=50%	Impute with similar/closest bad rate group
			If bad rate is very different from all other categories, then assign a special value and include in the regression analysis
		>50%	Create a dummy / indicator variable with missing (as 1) vs. non missing category
	>25	<=10%	Impute with the mean / median value
		>10% & <=50%	Assign a special value to the missing category in the original variable and create an indicator variable with missing value as 1 and others as 0
		>50%	Create a dummy / indicator variable with missing value as 1 and others as 0

SQL

SQL is a standard language for accessing and manipulating databases.

What is SQL?

- SQL stands for Structured Query Language
- SQL lets you access and manipulate databases
- SQL is an ANSI (American National Standards Institute) standard

What Can SQL do?

- SQL can execute queries against a database
- SQL can retrieve data from a database
- SQL can insert records in a database
- SQL can update records in a database
- SQL can delete records from a database
- SQL can create new databases
- SQL can create new tables in a database
- SQL can create stored procedures in a database
- SQL can create views in a database
- SQL can set permissions on tables, procedures, and views

Structured Query Language (SQL) is the relational database model's standard language.

SQL enable users to

- Create database and table structure (Data definition)
- Perform various types of data manipulation
- Query the database to extract useful information

Some Common SQL Data Types

Data Type	Format
Numeric	NUMBER(L,D), INTEGER,SMALLINT,DECIMAL(L,D)
Character	CHAR(L),VARCHAR(L)
Date	DATE

Listing the Table Contents with No Condition (s)

List all data in product table (* in SQL means all attributes (fields) in a table)

SELECT * FROM PRODUCT;

Choosing columns (Instead of using *, you may include all or part field names)

SELECT P_CODE, P_DESCRIPT, P_INDATE, P_ONHAND, P_MIN, P-PRICE, P_DISCOUNT, V_CODE FROM PRODUCT;

The Contents of the PRODUCT Table

P_CODE	P_DESCRIPT	P_INDATE	P_ONHAND	P_MIN	P_PRICE
1QER/31	Power painter, 15 psi., 3-nozzle	Friday, July 02, 1999	8	5	$109.99
13-Q2/P2	7.25-in. pwr. saw blade	Thursday, August 12, 1999	32	15	$14.99
14-Q1/L3	9.00-in. pwr. saw blade	Thursday, August 12, 1999	18	12	$17.49
1546-QQ2	Hrd. cloth, 1/4-in., 2x50	Wednesday, July 14, 1999	15	8	$39.95
1558-QW1	Hrd. cloth, 1/2-in., 3x50	Saturday, August 14, 1999	23	5	$43.99
2232/QTY	B&D jigsaw, 12-in. blade	Thursday, July 29, 1999	8	5	$109.92
2232/QWE	B&D jigsaw, 8-in. blade	Monday, August 23, 1999	6	5	$99.87
2238/QPD	B&D cordless drill, 1/2-in.	Saturday, June 19, 1999	12	5	$38.95
23109-HB	Claw hammer	Thursday, August 19, 1999	23	10	$5.95
23114-AA	Sledge hammer, 12 lb.	Wednesday, September 01, 1999	8	5	$14.40
54778-2T	Rat-tail file, 1/8-in. fine	Saturday, August 14, 1999	43	20	$4.99
89-WRE-Q	Hicut chain saw, 16 in.	Tuesday, July 06, 1999	11	5	$256.99
PVC23DRT	PVC pipe, 3.5-in., 8-ft	Thursday, August 19, 1999	188	75	$5.87
SM-18277	1.25-in. metal screw, 25	Wednesday, July 28, 1999	172	75	$6.99
SW-23116	2.5-in. wd. screw, 50	Friday, July 23, 1999	237	100	$8.45
WR3/TT3	Steel matting, 4'x8'x1/6", .5" mesh	Monday, August 16, 1999	18	5	$119.95

P_DISCOUNT	V_CODE
0.00	25595
0.05	21344
0.00	21344
0.00	23119
0.00	23119
0.05	24288
0.05	24288
0.05	25595
0.10	21225
0.05	
0.00	21344
0.05	24288
0.00	
0.00	21225
0.00	21231
0.10	25595

PRODUCT query results, continued →

Queries

Partial Listing of Table Contents with Condition (s)

Syntax
SELECT <column(s)>
FROM <table name>
WHERE <conditions>;

An example:
SELECT P_DESCRIPT, P_INDATE, P_PRICE, V_CODE
FROM PRODUCT
WHERE V_CODE = 21344;

P_DESCRIPT	P_INDATE	P_PRICE	V_CODE
7.25-in. pwr. saw blade	Thursday, August 12, 1999	$14.99	21344
9.00-in. pwr. saw blade	Thursday, August 12, 1999	$17.49	21344
Rat-tail file, 1/8-in. fine	Saturday, August 14, 1999	$4.99	21344

FIGURE 3.5 ■ SELECTED PRODUCT TABLE ATTRIBUTES FOR VENDOR CODE 21344

Specifying condition (s) with character or numeric fields

SELECT P_DESCRIPT, P_INDATE, P_PRICE, V_CODE
FROM PRODUCT
WHERE V_CODE = 21344;

The above SQL is correct if v_code (vendor_code) is defined as number.

If V-Code is defined as character the correct SQL statement is:

SELECT P_DESCRIPT, P_INDATE, P_PRICE, V_CODE
FROM PRODUCT
WHERE V_CODE = '21344';

Notice that the use of single quotation to quote v_code.

The WHERE clause can contain several conditions linked by AND or OR.

Select *
        from customer
                where state ='MO'
                        and city='Cape Girardeau'

This query lists all customers who lives in Cape Girardeau, MO. Notice that since state and city is character field, we use quotation.

Select * from friend
                where height > 5
                and weight <150

This query lists all friends whose height is taller than 5 feet and whose weight is less than 150 pounds. Notice that since height and weight are numerical fields, no quotation is necessary.

Ordering a Listing or Sorting

ORDER BY <attributes>

SELECT P_CODE, P_DESCRIPT, P_INDATE, P_PRICE
FROM PRODUCT
ORDER BY P_PRICE;

The above SQL lists output order by P_price in an ascending order

ORDER BY <attributes> desc

SELECT P_CODE, P_DESCRIPT, P_INDATE, P_PRICE
FROM PRODUCT
ORDER BY P_PRICE desc;

This SQL lists output order by P_price in a descending order

Some Basic SQL Numeric Functions

FUNCTION	OUTPUT
COUNT	The number of rows containing the specified attribute.
MIN	The minimum attribute value encountered.
MAX	The maximum attribute value encountered.
AVG	The arithmetic mean (average) for the specified attribute.
SUM	The sum of all values for a selected attribute.

Grouping

SELECT ...... FROM ...... WHERE condition ;
GROUP BY groupexpr [HAVING requirement]

List the number of students of each class.

SELECT class, COUNT(*) FROM student ;
                    GROUP BY class

class	cnt
1A	10
1B	9
1C	9
2A	8
2B	8
2C	6

List the Average Math test score of each class.

SELECT class, AVG (mtest) FROM student ;
GROUP BY class

class	avg_mtest
1A	85.90
1B	70.33
1C	37.89
2A	89.38
2B	53.13
2C	32.67

List the number of girls of each district

SELECT dcode, COUNT(*) FROM student ;
WHERE sex="F"
 GROUP BY dcode

dcode	cnt
HHM	6
KWC	1
MKK	1
SSP	5
TST	4
YMT	8

Sql Joins
Inner joins

Inner joins are the most common join type. An inner join simply looks for two rows that put together satisfy a join predicate. For example, this query uses the join predicate "S.Cust_Id = C.Cust_Id" to find all Sales and Customer rows with the same Cust_Id:

Select *
from Sales S inner join Customers C
on S.Cust_Id = C.Cust_Id

```
Cust_Id     Item       Cust_Id     Cust_Name
----------- ---------- ----------- ----------
2           Camera     2           John Doe
3           Computer   3           Jane Doe
3           Monitor    3           Jane Doe
```

Cust_Id 3 bought two items so this customer row appears twice in the result.

Cust_Id 1 did not purchase anything and so does not appear in the result.

We sold a 'Printer' to Cust_Id 4. There is no such customer so this sale does not appear in the result.

Inner joins are fully commutative. "A inner join B" and "B inner join A" are equivalent.

Outer joins

Suppose that we would like to see a list of all sales; even those that do not have a matching customer. We can write this query using an outer join. An outer join preserves all rows in one or both of the input tables even if we cannot find a matching row per the join predicate. For example:

select *
from Sales S left outer join Customers C
on S.Cust_Id = C.Cust_Id

```
Cust Id     Item       Cust Id     Cust Name
----------- ---------- ----------- ----------
2           Camera     2           John Doe
3           Computer   3           Jane Doe
3           Monitor    3           Jane Doe
4           Printer    NULL        NULL
```

Note that the server returns NULLs for the customer data associated with the 'Printer' sale since there is no matching customer. We refer to this row as "NULL extended."

Using a full outer join, we can find all customers regardless of whether they purchased anything and all sales regardless of whether they have a valid customer:

select *
from Sales S full outer join Customers C
on S.Cust_Id = C.Cust_Id

```
Cust Id      Item        Cust Id      Cust Name
-----------  ----------  -----------  ----------
2            Camera      2            John Doe
3            Computer    3            Jane Doe
3            Monitor     3            Jane Doe
4            Printer     NULL         NULL
NULL         NULL        1            Craig
```

The following table shows which rows will be preserved or NULL extended for each outer join variation:

Join	Preserve …
A left outer join B	all A rows
A right outer join B	all B rows
A full outer join B	all A and B rows

Full outer joins are commutative. In addition, "A left outer join B" and "B right outer join A" are equivalent.

Cross joins

A cross join performs a full Cartesian product of two tables. That is, it matches every row of one table with every row of another table. You cannot specify a join predicate for a cross join using the ON clause though you can use a WHERE clause to achieve essentially the same result as an inner join.

Cross joins are fairly uncommon. Two large tables should never be cross joined as this will result in a very expensive operation and a very large result set.
Select *from Sales S cross join Customers C

```
Cust Id      Item        Cust Id      Cust Name
-----------  ----------  -----------  ----------
2            Camera      1            Craig
3            Computer    1            Craig
3            Monitor     1            Craig
4            Printer     1            Craig
2            Camera      2            John Doe
3            Computer    2            John Doe
3            Monitor     2            John Doe
4            Printer     2            John Doe
2            Camera      3            Jane Doe
3            Computer    3            Jane Doe
3            Monitor     3            Jane Doe
4            Printer     3            Jane Doe
```

Model Building

1. Checking Avaibility of observations
2. Descriptive Statistics
3. Outlier Treatment
4. Missing value Treatment
5. Create New variable
6. Variable Analysis and Reduction
7. Splitting the dataset into development and validation
8. Model building
9. Analysis

Checking Avaibility of observations

- Total number of observations
- Number of available observations
- Percentage of Available observation
- Number of missing observations
- Percentage of missing observations
- Number of positive values
- Number of negative Values
- Number of observations with zero value

Descriptive Statistics

- Minimum value
- Maximum value
- Mean
- Median
- Standard deviation
- Skewness
- Kurtosis

Outlier Treatment

How to Detect Outliers
- Dot plot or Scatter plot
- Box plot

What Should We Do About Them?
- Transformation
- Deletion
- Winsorised Mean
- Trimmed Mean

Missing value Treatment

- Dropping variables (when more than 70 % of the data are missing, if the variable is very important we will go upto 60%)
- Listwise /Casewise deletion
- Business Ratios and using ranges to cap/delete observation
- Nominal variables: Treat missing data as just another category
- Substituted (plugged in) values, i.e. (Single) Imputation
  1. Mean
  2. Subgroup Mean
  3. Median
  4. Subgroup Median
  5. A regression estimate
- Maximum Likelihood Estimation and Multiple Imputation

Variable Analysis and Reduction

- Chi-squared statistic
- Information value
- Spearman rank order correlation coefficient
- Clustering techniques
- Multicollinearity check
  1. Correlation Matrix (R value must lie between -0.4 and 0.4, ideally)

  2. Tolerance (should be more than 0.4, ideally)

  3. Variance Inflation Factor (should be less than 2.5, ideally)

Splitting the Dataset into Development and Validation Sample

Development Sample (80% of the full file)

Validation sample (20% of the full file)

Situation may change if the number of observation in the full dataset is small

Regression Modelling

1. Transformation of variables (both dependent and independent variables, if required)
2. Transforming back to get original values (if dependent variable is transformed)
3. Create dummy variables for categorical variables

 Check list

- F-test (the relevant p-value must be less than 0.05).
- R-square and Adjusted R^2 (should not be less than 0.5, the more the better)
- Tolerance (should be more than 0.4, ideally)

- VIF (should be less than 2.5, ideally)
- T-test for final model variables (the relevant p-value must be less than 0.05)

Logistic Modelling

Transformation of variables (independent variables, if required)

Create dummy variables for categorical variables

Check list

- Concordant (should be between 65% and 85%)
- Discordant (should be between 15% and 35%)
- KS should not be beyond 3rd decile
- Rank ordering of bads
- Model must hold in k-cross validation samples
- No flipping of signs across samples
- Beta coeff should be close in dev and val samples
- Validation KS should be less than Dev KS by max 2%
- Variables should have logical trends
- Somer's D, Gamma, Tau-a, C (Should vary between 0 and 1, with larger values corresponding to stronger association between predicted and observed values)
- Hosmer-Lemeshow Goodness of Fit Test (Large values of χ^2_{HL} (and small *p*-values) indicate a lack of fit of the model)

Analysis

- Rank ordering
- KS Statistic (should be between 40 and 70)
- Lift Curve
- Divergence

Validation Check

1. Check whether the same variables come in the model significantly (p-value)
2. Check whether the same variables come in the model with same sign
3. KS Statistic (should be between 40 and 70)
4. Lift Curve
5. Divergence

Information Value:

$$\sum_{i-1}^{n} (\% good_i - \% bad_i) * \ln(\frac{\% G_i}{\% B_i})$$

IV = woe*(%good - %bad)

- **<0.02 Not predictive**
- **0.02- 0.1 Weak**
- **0.1-0.3 Medium**

- **>0.3 Strong**

Note: Variable with IV > 0.5 should be investigated for over predicting.

a.　　**Weight of Evidence**

Another consideration when defining the groups during formatting is the Weight of Evidence (WOE). The WOE measures the strength of each attribute or grouped attributes, in separating the 'good' and 'bad' accounts. It is a measure of the difference between the proportion of 'goods' and 'bads' in each attribute. The WOE is based on the log of odds calculation:

$$WOE = \ln \left(\frac{\text{Goods / Total Goods}}{\text{Bads / Total Bads}} \right) \times 100$$

i.e.

WOE = ln(%good/%bad)

HR

General Guidelines in Answering Interview Questions

Everyone is nervous on interviews. If you simply allow yourself to feel nervous, you'll do much better. Remember also that it's difficult for the interviewer as well. In general, be upbeat and positive. Never be negative. Rehearse your answers and time them. Never talk for more than 2 minutes straight.

Don't try to memorize answers word for word. Use the answers shown here as a guide only, and don't be afraid to include your own thoughts and words. To help you remember key concepts, jot down and review a few key words for each answer. Rehearse your answers frequently, and they will come to you naturally in interviews.

Tell me about yourself.

TRAPS: Beware, about 80% of all interviews begin with this "innocent" question. Many candidates, unprepared for the question, skewer themselves by rambling, recapping their life story, delving into ancient work history or personal matters.

BEST ANSWER: Start with the present and tell why you are well qualified for the position. Remember that the key to all successful interviewing is to match your qualifications to what the interviewer is looking for. In other words you must sell what the buyer is buying. This is the single most important strategy in job hunting.

So, before you answer this or any question it's imperative that you try to uncover your interviewer's greatest need, want, problem or goal.

To do so, make you take these two steps:

Do all the homework you can before the interview to uncover this person's wants and needs (not the generalized needs of the industry or company)

As early as you can in the interview, ask for a more complete description of what the position entails. You might say: "I have a number of accomplishments I'd like to tell you about, but I want to make the best use of our time together and talk directly to your needs. To help me do, that, could you tell me more about the most important priorities of this position? All I know is what I (heard from the recruiter, read in the classified ad, etc.)"

Then, ALWAYS follow-up with a second and possibly, third question, to draw out his needs even more. Surprisingly, it's usually this second or third question that unearths what the interviewer is most looking for.

You might ask simply, "And in addition to that?..." or, "Is there anything else you see as essential to success in this position?:

This process will not feel easy or natural at first, because it is easier simply to answer questions, but only if you uncover the employer's wants and needs will your answers make the most sense. Practice asking these key questions before giving your answers, the process will feel more natural and you will be light years ahead of the other job candidates you're competing with.

After uncovering what the employer is looking for, describe why the needs of this job bear striking parallels to tasks you've succeeded at before. Be sure to illustrate with specific examples of your

responsibilities and especially your achievements, all of which are geared to present yourself as a perfect match for the needs he has just described.

What are your greatest strengths?

TRAPS: This question seems like a softball lob, but be prepared. You don't want to come across as egotistical or arrogant. Neither is this a time to be humble.

BEST ANSWER: You know that your key strategy is to first uncover your interviewer's greatest wants and needs before you answer questions. And from Question 1, you know how to do this.

Prior to any interview, you should have a list mentally prepared of your greatest strengths. You should also have, a specific example or two, which illustrates each strength, an example chosen from your most recent and most impressive achievements.

You should, have this list of your greatest strengths and corresponding examples from your achievements so well committed to memory that you can recite them cold after being shaken awake at 2:30AM.

Then, once you uncover your interviewer's greatest wants and needs, you can choose those achievements from your list that best match up.

As a general guideline, the 10 most desirable traits that all employers love to see in their employees are:

A proven track record as an achiever...especially if your achievements match up with the employer's greatest wants and needs.

Intelligence...management "savvy".

Honesty...integrity...a decent human being.

Good fit with corporate culture...someone to feel comfortable with...a team player who meshes well with interviewer's team.

Likeability...positive attitude...sense of humor.

Good communication skills.

Dedication...willingness to walk the extra mile to achieve excellence.

Definiteness of purpose...clear goals.

Enthusiasm...high level of motivation.

Confident...healthy...a leader.

What are your greatest weaknesses?

TRAPS: Beware - this is an eliminator question, designed to shorten the candidate list. Any admission of a weakness or fault will earn you an "A" for honesty, but an "F" for the interview.

PASSABLE ANSWER: Disguise a strength as a weakness.

Example: "I sometimes push my people too hard. I like to work with a sense of urgency and everyone is not always on the same wavelength."

Drawback: This strategy is better than admitting a flaw, but it's so widely used, it is transparent to any experienced interviewer.

BEST ANSWER: (and another reason it's so important to get a thorough description of your interviewer's needs before you answer questions): Assure the interviewer that you can think of nothing that would

stand in the way of your performing in this position with excellence. Then, quickly review you strongest qualifications.

Example: "Nobody's perfect, but based on what you've told me about this position, I believe I' d make an outstanding match. I know that when I hire people, I look for two things most of all. Do they have the qualifications to do the job well, and the motivation to do it well? Everything in my background shows I have both the qualifications and a strong desire to achieve excellence in whatever I take on. So I can say in all honesty that I see nothing that would cause you even a small concern about my ability or my strong desire to perform this job with excellence."

Alternate strategy (if you don't yet know enough about the position to talk about such a perfect fit): Instead of confessing a weakness, describe what you like most and like least, making sure that what you like most matches up with the most important qualification for success in the position, and what you like least is not essential.

Example: Let's say you're applying for a teaching position. "If given a choice, I like to spend as much time as possible in front of my prospects selling, as opposed to shuffling paperwork back at the office. Of course, I long ago learned the importance of filing paperwork properly, and I do it conscientiously. But what I really love to do is sell (if your interviewer were a sales manager, this should be music to his ears.)

Tell me about something you did – or failed to do – that you now feel a little ashamed of.

TRAPS: There are some questions your interviewer has no business asking, and this is one. But while you may feel like answering, "none of your business," naturally you can't. Some interviewers ask this question on the chance you admit to something, but if not, at least they'll see how you think on your feet.

Some unprepared candidates, flustered by this question, unburden themselves of guilt from their personal life or career, perhaps expressing regrets regarding a parent, spouse, child, etc. All such answers can be disastrous.

BEST ANSWER: As with faults and weaknesses, never confess a regret. But don't seem as if you're stonewalling either.

Best strategy: Say you harbor no regrets, then add a principle or habit you practice regularly for healthy human relations.

Example: Pause for reflection, as if the question never occurred to you. Then say, "You know, I really can't think of anything." (Pause again, then add): "I would add that as a general management principle, I've found that the best way to avoid regrets is to avoid causing them in the first place. I practice one habit that helps me a great deal in this regard. At the end of each day, I mentally review the day's events and conversations to take a second look at the people and developments I'm involved with and do a doublecheck of what they're likely to be feeling. Sometimes I'll see things that do need more follow-up, whether a pat on the back, or maybe a five minute chat in someone's office to make sure we're clear on things…whatever."

"I also like to make each person feel like a member of an elite team, like the Boston Celtics or LA Lakers in their prime. I've found that if you let each team member know you expect excellence in their performance…if you work hard to set an example yourself…and if you let people know you appreciate and

respect their feelings, you wind up with a highly motivated group, a team that's having fun at work because they're striving for excellence rather than brooding over slights or regrets."

Why are you leaving (or did you leave) this position?

TRAPS: Never badmouth your previous industry, company, board, boss, staff, employees or customers. This rule is inviolable: never be negative. Any mud you hurl will only soil your suit.

Especially avoid words like "personality clash", "didn't get along", or others which cast a shadow on your competence, integrity, or temperament.

BEST ANSWER:

(If you have a job presently) If you're not yet 100% committed to leaving your present post, don't be afraid to say so. Since you have a job, you are in a stronger position than someone who does not. But don't be coy either. State honestly what you'd be hoping to find in a new spot. Of course, as stated often before, you answer will all the stronger if you have already uncovered what this position is all about and you match your desires to it.

(If you do not presently have a job.) Never lie about having been fired. It's unethical – and too easily checked. But do try to deflect the reason from you personally. If your firing was the result of a takeover, merger, division wide layoff, etc., so much the better.

But you should also do something totally unnatural that will demonstrate consummate professionalism. Even if it hurts , describe your own firing – candidly, succinctly and without a trace of bitterness – from the company's point-of-view, indicating that you could understand why it happened and you might have made the same decision yourself.

Your stature will rise immensely and, most important of all, you will show you are healed from the wounds inflicted by the firing. You will enhance your image as first-class management material and stand head and shoulders above the legions of firing victims who, at the slightest

provocation, zip open their shirts to expose their battle scars and decry the unfairness of it all.

For all prior positions: Make sure you've prepared a brief reason for leaving. Best reasons: more money, opportunity, responsibility or growth.

The "Silent Treatment"

TRAPS: Beware – if you are unprepared for this question, you will probably not handle it right and possibly blow the interview. Thank goodness most interviewers don't employ it. It's normally used by those determined to see how you respond under stress. Here's how it works:

You answer an interviewer's question and then, instead of asking another, he just stares at you in a deafening silence.

You wait, growing a bit uneasy, and there he sits, silent as Mt. Rushmore, as if he doesn't believe what you've just said, or perhaps making you feel that you've unwittingly violated some cardinal rule of interview etiquette.

When you get this silent treatment after answering a particularly difficult question , such as "tell me about your weaknesses", its intimidating effect can be most disquieting, even to polished job hunters.

Most unprepared candidates rush in to fill the void of silence, viewing prolonged, uncomfortable silences as an invitation to clear up the previous answer which has obviously caused some problem. And that's what they do – ramble on, sputtering more and more information, sometimes irrelevant and often damaging, because they are suddenly playing the role of someone who's goofed and is now trying to recoup. But since the candidate doesn't know where or how he goofed, he just keeps talking, showing how flustered and confused he is by the interviewer's unmovable silence.

BEST ANSWER: Like a primitive tribal mask, the Silent Treatment loses all it power to frighten you once you refuse to be intimidated. If your interviewer pulls it, keep quiet yourself for a while and then ask, with sincere politeness and not a trace of sarcasm, "Is there anything else I can fill in on that point?" That's all there is to it.

Whatever you do, don't let the Silent Treatment intimidate you into talking a blue streak, because you could easily talk yourself out of the position.

Why should I hire you?

TRAPS: Believe it or not, this is a killer question because so many candidates are unprepared for it. If you stammer or adlib you've blown it.

BEST ANSWER: By now you can see how critical it is to apply the overall strategy of uncovering the employer's needs before you answer questions. If you know the employer's greatest needs and desires, this question will give you a big leg up over other candidates because you will give him better reasons for hiring you than anyone else is likely to…reasons tied directly to his needs.

Whether your interviewer asks you this question explicitly or not, this is the most important question of your interview because he must answer this question favorably in is own mind before you will be hired. So help him out! Walk through each of the position's requirements as you understand them, and follow each with a reason why you meet that requirement so well.

Example: "As I understand your needs, you are first and foremost looking for someone who can manage the sales and marketing of your book publishing division. As you've said you need someone with a strong

background in trade book sales. This is where I've spent almost all of my career, so I've chalked up 18 years of experience exactly in this area. I believe that I know the right contacts, methods, principles, and successful management techniques as well as any person can in our industry."

"You also need someone who can expand your book distribution channels. In my prior post, my innovative promotional ideas doubled, then tripled, the number of outlets selling our books. I'm confident I can do the same for you."

"You need someone to give a new shot in the arm to your mail order sales, someone who knows how to sell in space and direct mail media. Here, too, I believe I have exactly the experience you need. In the last five years, I've increased our mail order book sales from $600,000 to $2,800,000, and now we're the country's second leading marketer of scientific and medical books by mail." Etc., etc., etc.,

Every one of these selling "couplets" (his need matched by your qualifications) is a touchdown that runs up your score. IT is your best opportunity to outsell your competition.

Where do you see yourself five years from now?

TRAPS: One reason interviewers ask this question is to see if you're settling for this position, using it merely as a stopover until something better comes along. Or they could be trying to gauge your level of ambition. If you're too specific, i.e., naming the promotions you someday hope to win, you'll sound presumptuous. If you're too vague, you'll seem rudderless.

BEST ANSWER: Reassure your interviewer that you're looking to make a long-term commitment…that this position entails exactly what you're looking to do and what you do extremely well. As for your future, you believe that if you perform each job at hand with excellence, future opportunities will take care of themselves.

Example: "I am definitely interested in making a long-term commitment to my next position. Judging by what you've told me about this position, it's exactly what I'm looking for and what I am very well qualified to do. In terms of my future career path, I'm confident that if I do my work with excellence, opportunities will inevitable open up for me. It's always been that way in my career, and I'm confident I'll have similar opportunities here."

Describe your ideal company, location and job.

TRAPS: This is often asked by an experienced interviewer who thinks you may be overqualified, but knows better than to show his hand by posing his objection directly. So he'll use this question instead, which often gets a candidate to reveal that, indeed, he or she is looking for something other than the position at hand.

BEST ANSWER: The only right answer is to describe what this company is offering, being sure to make your answer believable with specific reasons, stated with sincerity, why each quality represented by this opportunity is attractive to you.

Remember that if you're coming from a company that's the leader in its field or from a glamorous or much admired company, industry, city or position, your interviewer and his company may well have an "Avis" complex. That is, they may feel a bit defensive about being "second best" to the place you're coming from, worried that you may consider them bush league.

This anxiety could well be there even though you've done nothing to inspire it. You must go out of your way to assuage such anxiety, even if it's not expressed, by putting their virtues high on the list of exactly what you're looking for, providing credible reason for wanting these qualities.

If you do not express genuine enthusiasm for the firm, its culture, location, industry, etc., you may fail to answer this "Avis" complex objection and, as a result, leave the interviewer suspecting that a hot shot like you, coming from a Fortune 500 company in New York, just wouldn't be happy at an unknown manufacturer based in Topeka, Kansas.

Why do you want to work at our company?

TRAPS: This question tests whether you've done any homework about the firm. If you haven't, you lose. If you have, you win big.

BEST ANSWER: This question is your opportunity to hit the ball out of the park, thanks to the in-depth research you should do before any interview.

Best sources for researching your target company: annual reports, the corporate newsletter, contacts you know at the company or its suppliers, advertisements, articles about the company in the trade press.

What are your career options right now?

TRAPS: The interviewer is trying to find out, "How desperate are you?"

BEST ANSWER: Prepare for this question by thinking of how you can position yourself as a desired commodity. If you are still working, describe the possibilities at your present firm and why, though you're greatly appreciated there, you're looking for something more (challenge, money, responsibility, etc.). Also mention that you're seriously exploring opportunities with one or two other firms.

If you're not working, you can talk about other employment possibilities you're actually exploring. But do this with a light touch, speaking only in general terms. You don't want to seem manipulative or coy.

Why have you been out of work so long?

TRAPS: A tough question if you've been on the beach a long time. You don't want to seem like damaged goods.

BEST ANSWER: You want to emphasize factors which have prolonged your job search by your own choice. Example: "After my job was terminated, I made a conscious decision not to jump on the first opportunities to come along. In my life, I've found out that you can always turn a negative into a positive IF you try hard enough. This is what I determined to do. I decided to take whatever time I needed to think through what I do best, what I most want to do, where I'd like to do it…and then identify those companies that could offer such an opportunity."

"Also, in all honesty, you have to factor in the recession (consolidation, stabilization, etc.) in the (banking, financial services, manufacturing, advertising, etc.) industry."

"So between my being selective and the companies in our industry downsizing, the process has taken time. But in the end, I'm convinced that when I do find the right match, all that careful evaluation from both sides of the desk will have been well worthwhile for both the company that hires me and myself.

Tell me honestly about the strong points and weak points of your boss (company, management team, etc.)…

TRAPS: Skillfull interviewers sometimes make it almost irresistible to open up and air a little dirty laundry from your previous position. DON'T

BEST ANSWER: Remember the rule: Never be negative. Stress only the good points, no matter how charmingly you're invited to be critical.

Your interviewer doesn't care a whit about your previous boss. He wants to find out how loyal and positive you are, and whether you'll criticize him behind his back if pressed to do so by someone in this own company. This question is your opportunity to demonstrate your loyalty to those you work with.

What are your outside interests?

TRAPS: You want to be a well-rounded, not a drone. But your potential employer would be even more turned off if he suspects that your heavy extracurricular load will interfere with your commitment to your work duties.

BEST ANSWERS: Try to gauge how this company's culture would look upon your favorite outside activities and be guided accordingly.

You can also use this question to shatter any stereotypes that could limit your chances. If you're over 50, for example, describe your activities that demonstrate physical stamina. If you're young, mention an activity that connotes wisdom and institutional trust, such as serving on the board of a popular charity.

But above all, remember that your employer is hiring your for what you can do for him, not your family, yourself or outside organizations, no matter how admirable those activities may be.

On confidential matters…

TRAPS: When an interviewer presses you to reveal confidential information about a present or former employer, you may feel it's a no-win situation. If you cooperate, you could be judged untrustworthy. If you don't, you may irritate the interviewer and seem obstinate, uncooperative or overly suspicious.

BEST ANSWER: Your interviewer may press you for this information for two reasons.

First, many companies use interviews to research the competition. It's a perfect set-up. Here in their own lair, is an insider from the enemy camp who can reveal prized information on the competition's plans, research, financial condition, etc.

Second, the company may be testing your integrity to see if you can be cajoled or bullied into revealing confidential data.

What to do? The answer here is easy. Never reveal anything truly confidential about a present or former employer. By all means, explain your reticence diplomatically. For example, "I certainly want to be as open as I can about that. But I also wish to respect the rights of those who have trusted me with their most sensitive information, just as you would hope to be able to trust any of your key people when talking with a competitor…"

And certainly you can allude to your finest achievements in specific ways that don't reveal the combination to the company safe.

But be guided by the golden rule. If you were the owner of your present company, would you feel it ethically wrong for the information to be given to your competitors? If so, steadfastly refuse to reveal it.

Remember that this question pits your desire to be cooperative against your integrity. Faced with any such choice, always choose integrity. It is a far more valuable commodity than whatever information the company may pry from you. Moreover, once you surrender the information, your stock goes down. They will surely lose respect for you.

One President we know always presses candidates unmercifully for confidential information. If he doesn't get it, he grows visibly annoyed, relentlessly inquisitive, It's all an act. He couldn't care less about the information. This is his way of testing the candidate's moral fiber. Only those who hold fast are hired.

Looking back, what would you do differently in your life?

TRAPS: This question is usually asked to uncover any life-influencing mistakes, regrets, disappointments or problems that may continue to affect your personality and performance.

You do not want to give the interviewer anything negative to remember you by, such as some great personal or career disappointment, even long ago, that you wish could have been avoided.

Nor do you wish to give any answer which may hint that your whole heart and soul will not be in your work.

BEST ANSWER: Indicate that you are a happy, fulfilled, optimistic person and that, in general, you wouldn't change a thing.

Example: "It's been a good life, rich in learning and experience, and the best it yet to come. Every experience in life is a lesson it its own way. I wouldn't change a thing."

Could you have done better in your last job?

TRAPS: This is no time for true confessions of major or even minor problems.

BEST ANSWER: Again never be negative.

Example: "I suppose with the benefit of hindsight you can always find things to do better, of course, but off the top of my head, I can't think of anything of major consequence."

(If more explanation seems necessary) Describer a situation that didn't suffer because of you but from external conditions beyond your control.

For example, describe the disappointment you felt with a test campaign, new product launch, merger, etc., which looked promising at first, but led to underwhelming results. "I wish we could have known at the start what we later found out (about the economy turning, the marketplace changing, etc.), but since we couldn't, we just had to go for it. And we did learn from it…"

Question 24 Can you work under pressure?

TRAPS: An easy question, but you want to make your answer believable.

BEST ANSWER: Absolutely…(then prove it with a vivid example or two of a goal or project accomplished under severe pressure.)

What makes you angry?

TRAPS: You don't want to come across either as a hothead or a wimp.

BEST ANSWER: Give an answer that's suited to both your personality and the management style of the firm. Here, the homework you've done about the company and its style can help in your choice of words.

Examples: If you are a reserved person and/or the corporate culture is coolly professional:

"I'm an even-tempered and positive person by nature, and I believe this helps me a great deal in keeping my department running smoothly, harmoniously and with a genuine esprit de corps. I believe in communicating clearly what's expected, getting people's commitment to those goals, and then following up continuously to check progress."

"If anyone or anything is going off track, I want to know about it early. If, after that kind of open communication and follow up, someone isn't getting the job done, I'll want to know why. If there's no good reason, then I'll get impatient and angry…and take appropriate steps from there. But if you hire good people, motivate them to strive for excellence and then follow up constantly, it almost never gets to that state."

If you are feisty by nature and/or the position calls for a tough straw boss.

"You know what makes me angry? People who (the fill in the blanks with the most objectionable traits for this type of position)…people who don't pull their own weight, who are negative, people who lie…etc."

Question 26 Why aren't you earning more money at this stage of your career?

TRAPS: You don't want to give the impression that money is not important to you, yet you want to explain why your salary may be a little below industry standards.

BEST ANSWER: You like to make money, but other factors are even more important.

Example: "Making money is very important to me, and one reason I'm here is because I'm looking to make more. Throughout my career, what's been even more important to me is doing work I really like to do at the kind of company I like and respect.

(Then be prepared to be specific about what your ideal position and company would be like, matching them as closely as possible to the opportunity at hand.

Who has inspired you in your life and why?

TRAPS: The two traps here are unpreparedness and irrelevance. If you grope for an answer, it seems you've never been inspired. If you ramble about your high school basketball coach, you've wasted an opportunity to present qualities of great value to the company.

BEST ANSWER: Have a few heroes in mind, from your mental "Board of Directors" – Leaders in your industry, from history or anyone else who has been your mentor.

Be prepared to give examples of how their words, actions or teachings have helped inspire your achievements. As always, prepare an answer which highlights qualities that would be highly valuable in the position you are seeking.

What was the toughest decision you ever had to make?

TRAPS: Giving an unprepared or irrelevant answer.

BEST ANSWER: Be prepared with a good example, explaining why the decision was difficult…the process you followed in reaching it…the courageous or effective way you carried it out…and the beneficial results.

Tell me about the most boring job you've ever had.

TRAPS: You give a very memorable description of a very boring job. Result? You become associated with this boring job in the interviewer's mind.

BEST ANSWER: You have never allowed yourself to grow bored with a job and you can't understand it when others let themselves fall into that rut.

Example: "Perhaps I've been fortunate, but that I've never found myself bored with any job I have ever held. I've always enjoyed hard work. As with actors who feel there are no small parts, I also believe that in every company or department there are exciting challenges and intriguing problems crying out for energetic and enthusiastic solutions. If you're bored, it's probably because you're not challenging yourself to tackle those problems right under your nose."

Have you been absent from work more than a few days in any previous position?

TRAPS: If you've had a problem, you can't lie. You could easily be found out. Yet admitting an attendance problem could raise many flags.

BEST ANSWER: If you have had no problem, emphasize your excellent and consistent attendance record throughout your career.

Also describe how important you believe such consistent attendance is for a key executive…why it's up to you to set an example of dedication…and why there's just no substitute for being there with your people to keep the operation running smoothly, answer questions and handle problems and crises as they arise.

If you do have a past attendance problem, you want to minimize it, making it clear that it was an exceptional circumstance and that it's cause has been corrected.

To do this, give the same answer as above but preface it with something like, "Other that being out last year (or whenever) because of (your reason, which is now in the past), I have never had a problem and have enjoyed an excellent attendance record throughout my career.

Furthermore, I believe, consistent attendance is important because…" (Pick up the rest of the answer as outlined above.).

what changes would you make if you came on board?

TRAPS: Watch out! This question can derail your candidacy faster than a bomb on the tracks – and just as you are about to be hired.

Reason: No matter how bright you are, you cannot know the right actions to take in a position before you settle in and get to know the operation's strengths, weaknesses key people, financial condition, methods of operation, etc. If you lunge at this temptingly baited question, you will probably be seen as someone who shoots from the hip.

Moreover, no matter how comfortable you may feel with your interviewer, you are still an outsider. No one, including your interviewer, likes to think that a know-it-all outsider is going to come in, turn the place upside down and with sweeping, grand gestures, promptly demonstrate what jerks everybody's been for years.

BEST ANSWER: You, of course, will want to take a good hard look at everything the company is doing before making any recommendations.

Example: "Well, I wouldn't be a very good doctor if I gave my diagnosis before the examination. Should you hire me, as I hope you will, I'd want to take a good hard look at everything you're doing and understand why it's being done that way. I'd like to have in-depth meetings with you and the other key people to get a deeper grasp of what you feel you're doing right and what could be improved.

"From what you've told me so far, the areas of greatest concern to you are…" (name them. Then do two things. First, ask if these are in fact his major concerns. If so then reaffirm how your experience in meeting similar needs elsewhere might prove very helpful).

I'm concerned that you don't have as much experience as we'd like in…

TRAPS: This could be a make-or-break question. The interviewer mostly likes what he sees, but has doubts over one key area. If you can assure him on this point, the job may be yours.

BEST ANSWER: This question is related to "The Fatal Flaw" (Question 18), but here the concern is not that you are totally missing some qualifications, such as CPA certification, but rather that your experience is light in one area.

Before going into any interview, try to identify the weakest aspects of your candidacy from this company's point of view. Then prepare the best answer you possible can to shore up your defenses.

To get past this question with flying colors, you are going to rely on your master strategy of uncovering the employer's greatest wants and needs and then matching them with your strengths. Since you already know how to do this from Question 1, you are in a much stronger position.

More specifically, when the interviewer poses as objection like this, you should…

Agree on the importance of this qualification.

Explain that your strength may be indeed be greater than your resume indicates because…

When this strength is added to your other strengths, it's really your combination of qualifications that's most important.

Then review the areas of your greatest strengths that match up most favorably with the company's most urgently-felt wants and needs.

This is powerful way to handle this question for two reasons. First, you're giving your interviewer more ammunition in the area of his concern. But more importantly, you're shifting his focus away from this one, isolated area and putting it on the unique combination of

strengths you offer, strengths which tie in perfectly with his greatest wants.

How do you feel about working nights and weekends?

TRAPS: Blurt out "no way, Jose" and you can kiss the job offer goodbye. But what if you have a family and want to work a reasonably normal schedule? Is there a way to get both the job and the schedule you want?

BEST ANSWER: First, if you're a confirmed workaholic, this question is a softball lob. Whack it out of the park on the first swing by saying this kind of schedule is just your style. Add that your family understands it. Indeed, they're happy for you, as they know you get your greatest satisfaction from your work.

If however, you prefer a more balanced lifestyle, answer this question with another: "What's the norm for your best people here?"

If the hours still sound unrealistic for you, ask, "Do you have any top people who perform exceptionally for you, but who also have families and like to get home in time to see them at night?" Chances are this company does, and this associates you with this other "top-performers-who-leave-not-later-than-six" group.

Depending on the answer, be honest about how you would fit into the picture. If all those extra hours make you uncomfortable, say so, but phrase your response positively.

Example: "I love my work and do it exceptionally well. I think the results speak for themselves, especially in …(mention your two or three qualifications of greater interest to the employer. Remember, this is what he wants most, not a workaholic with weak credentials). Not only would I bring these qualities, but I've

built my whole career on working not just hard, but smart. I think you'll find me one of the most productive people here.

I do have a family who likes to see me after work and on weekends. They add balance and richness to my life, which in turn helps me be happy and productive at work. If I could handle some of the extra work at home in the evenings or on weekends, that would be ideal. You'd be getting a person of exceptional productivity who meets your needs with strong credentials. And I'd be able to handle some of the heavy workload at home where I can be under the same roof as my family. Everybody would win."

Are you willing to relocate or travel?

TRAPS: Answer with a flat "no" and you may slam the door shut on this opportunity. But what if you'd really prefer not to relocate or travel, yet wouldn't want to lose the job offer over it?

BEST ANSWER: First find out where you may have to relocate and how much travel may be involved. Then respond to the question.

If there's no problem, say so enthusiastically.

If you do have a reservation, there are two schools of thought on how to handle it.

One advises you to keep your options open and your reservations to yourself in the early going, by saying, "no problem". You strategy here is to get the best offer you can, then make a judgment whether it's worth it to you to relocate or travel.

Also, by the time the offer comes through, you may have other offers and can make a more informed decision. Why kill of this opportunity before it has chance to blossom into something really special? And if you're a little more desperate three months from now, you might wish you hadn't slammed the door on relocating or traveling.

The second way to handle this question is to voice a reservation, but assert that you'd be open to relocating (or traveling) for the right opportunity.

The answering strategy you choose depends on how eager you are for

My Sincere Thanks to

Sandeep Sunkara

Rami Reddy

V2K Vijay

Vineet

Gopal Prasad

Deepak Kumar

IIT Bombay ASI 06-08 Batch

IITB Placement team 2008