

Solutions Manual – Chapter 1

Solutions to Multiple Choice Questions

1. B
2. D
3. C
4. C
5. C
6. C
7. D
8. B
9. B
10. B

Solutions to Discussion Questions

1. Data analytics is defined as the process of evaluating data with the purpose of drawing conclusions to address business questions. Indeed, effective Data Analytics provides a way to search through large structured and unstructured data to identify unknown patterns or relationships.

A university might learn from the analyzing the demographics of its current set of students in order to attract its future student recruits. Did they come from cities or high schools that were close by? Were their parents alumni of the university? Did they score high on certain parts of the ACT? Were those offered a scholarship more likely to attend, etc.? Was social media effective in attracting students? By analyzing this type of data, previously unknown patterns will emerge that will make recruiting students more effective.

2. There are many potential answers. For example, Monsanto may use mathematical and statistical models to plot out the best times to plant both male and female plans and where to plant them to maximize yield. (https://www.cio.com/article/3221621/analytics/6-data-analytics-success-stories-an-inside-look.html#tk.cio_rs)
3. There are many potential answers. Accountants might use data analytics to learn more about their allowance for doubtful accounts by learning which customers pay or do not pay their receivable balances on a timely basis. This will help make a more accurate balance of net receivables.
4. There are many potential answers. For example, data analytics associated with financial reporting may help accountants determine if any of their inventory obsolete? It may also help the company benchmark on the financial statements and financial reporting of other similar companies and understand their accounting practices to help infer their own.
5. The impact cycle suggests an order of 1) Identifying the Questions; 2) Mastering the Data; 3) Performing the test plan; 4) Addressing and refining results; 5) Communicating insights and 6)

Tracking outcomes. The cycle starts with a question and then identifying data and test plan that might address that question. The results of the data analysis are communicated and tracked which may lead to additional, possibly more refined questions that then restart the cycle.

6. Data analysis is most effective when a question is identified that needs to be addressed. That will focus the analysis on which data and which test method might be most effective in addressing or answering the question.
7. Mastering the data requires one to know what data is available and whether it might be able to help address the business problem. We need to know everything about the data, including how to access it, its availability, how reliable it is (if there are errors), and what time periods it covers to make sure it coincides with the timing of our business problem, etc.
8. Alibaba uses the profiling data approach to identify potential cases of fraud. Alibaba has worked to capture fraud signals directly from its extensive database of user behaviors and its network, then analyzes them in real-time using machine learning to accurately sort the bad users from the good ones.
9. Facebook uses link prediction to predict a relationship between two people when it suggests people that one likely knows due to similar other friends, high schools, college or work locations, etc.
10. While sampling is useful, it is still just that, sampling. By looking at all of the transactions and testing them in a way that will highlight the ones that are the biggest dollar items, or are most unusual, that will allow auditors to focus on specific items that might be of material significance.
11. There are several correct answers. One data approach might be regression analysis where, given a balance of total accounts receivable held by a firm, how long it has been outstanding, if they have paid debts in the past all will help predict the appropriate level of allowance for doubtful accounts for bad debts.
12. The Debt-to-Income ratio might suggest to LendingClub that the person asking for the loan was simply asking for too big of a loan and they would have little ability to repay it. The lower the credit score, the less likely the loanee would be able to repay the loan.
13. There are many other potential predictors of whether the LendingClub would pay a loan. Here are a few possibilities: What other debt do they have? How much is their disposable income? Do they have a clean criminal record? Have they had a loan with LendingClub before and did they repay it?

Solutions to Problems

Note: Some problems and solutions may be altered in Connect for auto grading purposes.

Problem 1-1

Here are the predictive attributes and whether they would be applicable to predicting which loans would be delinquent and which loans will ultimately be fully repaid.

<u>Yes/No</u>	<u>Predictive Attributes</u>
No	desc (Loan description provided by borrower)
Yes	dti (Monthly debt payments to monthly income Ratio)
Yes	grade (LC assigned loan grade)
Yes	home_ownership (values include Rent, Own, Mortgage, Other)
No	next_pymnt_d (Next scheduled payment date)
No	term (The number of payments on the loan)
Yes	tot_cur_bal (Total current balance of all accounts)

Problem 1-2

Potential attributes from the RejectStats data dictionary that might help predict loan acceptance or rejection include the following:

- Amount Requested
- Risk_Score
- Debt-to-Income Ratio
- Zip Code
- State (Possibly)
- Employment Length

Problem 1-3

Percentage of total loans rejected that live in Arkansas = 1.219%

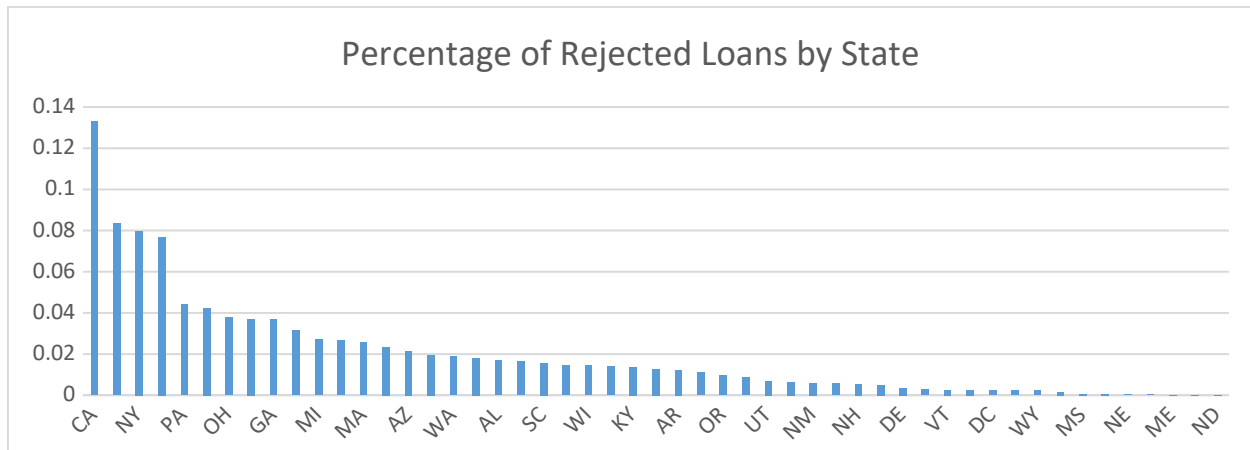
2,915,918 population in Arkansas divided by USA population of 308,745,538 = 0.9444%

The loan rejection percentage is greater than the percent of the USA population that lives in Arkansas (per 2010 census), but is reasonably close.

Problem 1-4

State	Loan Rejection %
CA	0.13292708
TX	0.08344411
NY	0.0797736
FL	0.07688089
PA	0.04401981
IL	0.04246422
OH	0.03779744
NJ	0.03708008
GA	0.03683527
VA	0.03131478
MI	0.02718255
NC	0.02672393
MA	0.02547822
MD	0.02340048
AZ	0.02142811
MO	0.01954559
WA	0.0187585
CO	0.01812325
AL	0.0169798
CT	0.01640652
SC	0.01569535
LA	0.01450077
WI	0.01430865
MN	0.01407314
KY	0.01367649
NV	0.01275305
AR	0.01219062
OK	0.01103943
OR	0.00954581
KS	0.00862547
UT	0.00692579
WV	0.00643153
NM	0.00590939
HI	0.005756
NH	0.00551739
RI	0.00498905
DE	0.00354346
MT	0.00284933
VT	0.00250537

AK	0.00249142
DC	0.00236128
SD	0.00223887
WY	0.00220479
IN	0.00149516
MS	0.00059962
TN	0.00055003
NE	0.00022311
IA	0.00017043
ME	0.0001379
ID	8.0568E-05
ND	4.6482E-05



The loan rejection percentage roughly corresponds with the population of each state. However, there is still substantial variation between the rejection percentage of each state.

Problem 1-5

Here is the pivot table by risk score grouping:

Row Labels	Count of Loan Title
Excellent	2931
Fair	236669
Good	83543
Poor	189621
Very Bad	145322
Very Good	11907
Grand Total	669993

The Excellent category had the smallest group, whereas the Fair group had the biggest group. Arguably there is a greater population of Fair, even though Very Bad has a smaller count, it is clearly the worst of the group.

Problem 1-6

Here is the pivot table by Debt-to-Income (DTI) grouping:

Row Labels	Count of Amount Requested
High	340862
Low	159464
Medium	169667
Grand Total	669993

Low DTI is the smallest grouping whereas High DTI has the largest grouping.

Problem 1-7

Here is the pivot table for the loans with excellent risks but high debt-to-incomes, by years of employment:

Row Labels	Count of Amount Requested
Excellent	2931
High	1190
0	942
1	12
2	14
3	11
4	12
5	92
6	9
7	15
8	9
9	6
10	68

Perhaps those with excellent credit just asked for too big of a loan given their existing debt and that is why they are rejected. This PivotTable analysis suggests those with excellent credit asked for a larger loan given the debt they already had as compared to any of the others, suggesting a reason why even those potential borrowers with excellent credit were rejected. There weren't a lot of them, but there were certainly some!