



# COMMERCIAL BANKING, CORP

COMMERCIAL BANKING, CORP

REQUEST FOR PROPOSAL

RFP #: IP – F1.H1

TITLE: BANKING INSURANCE PRODUCT – PHASE 1

CLOSING DATE AND TIME: SEPTEMBER 6, 2019 @ 5:00 PM

# Banking Insurance Product – Phase 1: IP – F1.H1

---

## Purpose

By responding to this Request for Proposal (RFP), the Proposer agrees that s/he has read and understood all documents within this RFP package.

## Submission Details

Responders to this RFP should supply:

- A business report **up to 4 pages** (not including cover page, table of contents, or any needed appendix), including any supporting plots and tables.
- The commented code used to produce the results.

The report should address **all points described in the “Objective” section** below.

The report should be returned in the following way:

- Electronic (mailto: [Aric\\_LaBarr@ncsu.edu](mailto:Aric_LaBarr@ncsu.edu); Subject Line: Banking Insurance Product – Phase 1)

## Background

The Commercial Banking Corporation (hereafter the “Bank”), acting by and through its department of *Customer Services and New Products* is seeking proposals for banking services. The Bank ultimately wants to predict which customers will buy a variable rate annuity product.

A variable annuity is a contract between you and an insurance company / bank, under which the insurer agrees to make periodic payments to you, beginning either immediately or at some future date. You purchase a variable annuity contract by making either a single purchase payment or a series of purchase payments.

A variable annuity offers a range of investment options. The value of your investment as a variable annuity owner will vary depending on the performance of the investment options you choose. The investment options for a variable annuity are typically mutual funds that invest in stocks, bonds, money market instruments, or some combination of the three. If you are interested in more information, see: <http://www.sec.gov/investor/pubs/varannty.htm>

The project will be broken down into 3 phases:

- Phase 1 – Variable Understanding and Assumptions
- Phase 2 – Variable Selection and Modeling Building
- Phase 3 – Model Assessment and Prediction

## Objective

The scope of services in this phase includes the following:

- For this phase use **only** the training data set.

- Explore the predictor variables **individually** with the target variable of whether the customer bought the insurance product.
  - Summarize **only the significant variables** in a table ranking from most significant to least significant – the Bank currently uses  $\alpha = 0.002$ , but is open to another if you defend your reason.
    - This table should separate out the four possible classes of variables – binary, ordinal, nominal, continuous.
    - (HINT: Explore the predictor variables individually for now since you have not yet accounted for missing values.)
  - In an appendix, include a table with **all** of the variables ranked by significance.
- Provide a table of odds ratios for **only binary predictor variables** in relation to the target variable.
  - Rank these odds ratios by magnitude.
  - Interpret **only** the highest magnitude odds ratio.
  - Report on any interesting findings.
    - (HINT: This is open-ended and has no correct answer. However, you should get use to keeping an eye out for what you might deem important or interesting when exploring data to report in an executive summary.)
- Provide a summary of results around the linearity assumption of continuous variables.
  - List both which variables meet and do not meet the needed assumption for continuous variables.
  - (HINT: Do not get overly mathematical here. Just report what you find; do not teach.)
- Provide a summary of important data considerations as follows:
  - Visual representation of which variables have the highest (defined by you for now) amount of missing values.
  - List any combinations of variables that you feel have redundant information so the Bank might consider removing them in the future.
    - (HINT: This is open-ended and has no correct answer. For example, presence of a money market account and money market balance.)
  - Report on any interesting findings.
    - (HINT: This is open-ended and has no correct answer. However, you should get use to keeping an eye out for what you might deem important or interesting when exploring data to report in an executive summary. For example, teller visits as well as other variables might represent human contact with the bank as compared to only online contact.)

## Data Provided

The following two sets of data are provided for the proposal:

- The training data set **insurance\_t** contains 8,495 observations and 48 variables.
  - All of these customers have been offered the product in the data set under the variable **INS**, which takes a value of 1 if they bought and 0 if they did not buy.
  - There are 47 variables describing the customer's attributes **before** they were offered the new insurance product.
- The validation data set **insurance\_v** contains 2,124 observations and 48 variables.
- The table below describes the Roles and Description of the variables found in both data sets.
  - **Except for Branch of Bank**, consider anything with more than 10 distinct values as continuous.
- (HINT: If you are using R, use the **haven** package and the **read\_sas()** function to open the **.sas7bdat** files.

<i>Name</i>	<i>Model Role</i>	<i>Description</i>
<i>ACCTAGE</i>	Input	Age of oldest account
<i>DDA</i>	Input	Indicator for checking account
<i>DDABAL</i>	Input	Checking account balance
<i>DEP</i>	Input	Checking deposits
<i>DEPAMT</i>	Input	Total amount deposited
<i>CASHBK</i>	Input	Number of cash back requests
<i>CHECKS</i>	Input	Number of checks written
<i>DIRDEP</i>	Input	Indicator for direct deposit
<i>NSF</i>	Input	Number of insufficient fund issues
<i>NSFAMT</i>	Input	Amount of NSF
<i>PHONE</i>	Input	Number of telephone banking interactions
<i>TELLER</i>	Input	Number of teller visit interactions
<i>SAV</i>	Input	Indicator for savings account
<i>SAVBAL</i>	Input	Savings account balance
<i>ATM</i>	Input	Indicator for ATM interaction
<i>ATMAMT</i>	Input	Total ATM withdrawal amount
<i>POS</i>	Input	Number of point of sale interactions
<i>POSAMT</i>	Input	Total amount for point of sale interactions
<i>CD</i>	Input	Indicator for certificate of deposit account
<i>CDBAL</i>	Input	CD balance
<i>IRA</i>	Input	Indicator for retirement account
<i>IRABAL</i>	Input	IRA balance
<i>LOC</i>	Input	Indicator for line of credit
<i>LOCBAL</i>	Input	LOC balance
<i>INV</i>	Input	Indicator for investment account
<i>INVBAL</i>	Input	INV balance
<i>ILS</i>	Input	Indicator for installment loan
<i>ILSBAL</i>	Input	ILS balance
<i>MM</i>	Input	Indicator for money market account
<i>MMBAL</i>	Input	MM balance
<i>MMCRED</i>	Input	Number of money market credits
<i>MTG</i>	Input	Indicator for mortgage
<i>MTGBAL</i>	Input	MTG balance
<i>CC</i>	Input	Indicator for credit card
<i>CCBAL</i>	Input	CC balance
<i>CCPURC</i>	Input	Number of credit card purchases
<i>SDB</i>	Input	Indicator for safety deposit box
<i>INCOME</i>	Input	Income
<i>HMOWN</i>	Input	Indicator for home ownership
<i>LORES</i>	Input	Length of residence in years
<i>HMVAL</i>	Input	Value of home

*AGE*  
*CRSCORE*  
*MOVED*  
*INAREA*  
*INS*  
*BRANCH*  
*RES*

Input	Age
Input	Credit score
Input	Recent address change
Input	Indicator for local address
Target	Indicator for purchase of insurance product
Input	Branch of bank
Input	Area classification