Auto Insurance Case Study

Thank you for your interest in joining the Data & Analytics team at Zurich North America. As part of the interview process, we have a case study for you to complete in 3 days.

We trust that you will complete this on your own and keep it confidential.

Thanks for taking the time to complete this case study.

#### The Data

Along with this document, you should have received two datasets:

1. **auto\_policies\_2017.csv**

This data set is a set of personal auto insurance policies taken out in 2017. There are 60,392 policies (rows), of which 10,030 had at least one claim.

1. **auto\_potential\_customers\_2018.csv**

This is a list of 7,464 potential customers for 2018.

A glossary for the datasets is provided below.

pol\_number policy number for the insurance policy

pol\_eff\_dt auto insurance policy effective date

gender gender of driver: F, M

agecat driver's age category: 1 (youngest), 2, 3, 4, 5, 6

date\_of\_birth driver's date of birth

credit\_score driver’s credit score(integer): 301-850, 301=excellent, 850=excellent

area driver's area of residence: A, B, C, D, E, F

traffic\_index traffic index of driver’s area of residence(integer): 100=country average, >100 means worse traffic conditions than average

veh\_age age of vehicle(categorical): 1 (youngest), 2, 3, 4

veh\_body vehicle body, coded as:

BUS

CONVT = convertible

COUPE

HBACK = hatchback

HDTOP = hardtop

MCARA = motorized caravan

MIBUS = minibus

PANVN = panel van

RDSTR = roadster

SEDAN

STNWG = station wagon

TRUCK

UTE = utility

veh\_value vehicle value, in $10,000s

claim\_office office location of claim handling agent: A, B, C, D

numclaims number of claims(integer): 0 if no claim

claimcst0 claim amount: 0 if no claim

annual\_premium total charged premium

#### The Case Study

The manager in charge of personal auto has asked you to help her:

1. target the potential customers in 2018 which will result in the lowest claim amounts. For example:
   1. find customers with a low probability of having any claims
   2. find customers with the lowest cost per claim, given that a claim occurs
2. create a marketing campaign that engages with potential 2018 customers based on their risk profiles (e.g. letting riskier customers know about educational opportunities, which will benefit both the customer and insurance company by reducing the risk of claims).  In order to do that, the manager must make several risk profile ‘groups’ to be used in this campaign.

You may answer this (these) questions using methodology and software of your choice.

#### Sharing Your Results

1. Presentation

Please create a 60 minute presentation:

* The first half should be targeted to the personal auto manager and to marketing. Please give them your recommendations, and explain how and why you are making those recommendations.
* The second half should be targeted to a group of data scientists. Please give your high level methodology and thought process.

1. **Please send your presentation, code and any additional analysis to Mathew Fish (mathew.fish @zurichna.com) and Diane Wagner (**[**diane.wagner@zurichna.com**](mailto:diane.wagner@zurichna.com)**) at the end of the third day.**

We will review and then let you know if we will bring you in to interview and present your case study.

Again, thanks for your interest in joining the team, and for taking the time to complete this case study. We look forward to seeing what you can do.