两个关于产品的销售推广的项目，包括大额存单目标客户定位、信用卡到保险基金的交叉销售模型。两个关于客户经营的项目，包括客户流失/存款流失预警和客户标签系统（客户画像）；

1. 大额存单目标客户定位模型：
2. 背景：花旗银行去年9月获得大额存单销售资格，产品上线后建立决策树模型定位大额存单目标客户群体；助力产品的销售推动；
3. 模型建立：选取全体借记卡客户作为样本（19万），过去六个月作为观察窗口，2017年10月作为表现窗口，已经购买大额存单的70位客户再加上购买了一年期以上，20万以上定期存款的客户作为目标变量；
4. 模型结论：选取决策树中responder rate最高的5片树叶作为目标客群；重要的特点包括：产品偏好（储蓄性质保险，债券，定存）；行为模式（过去3个月内开立过定期存款，未来一个月内有产品到期，过去半年曾登陆过网银或手机银行）；资产规模（AUM60万以上，活期存款余额2万美元12万人民币以上）
5. 模型表现：A.在训练数据集上的KS为43，提升图最大提升倍数为64倍数（最高responder rate 4%，随机responder rate 0.07%）; B.模型投入运行后，10月、11月、12月每个月选取500到1000个客户（19万/4的客户基础中选）发送给一线重点营销，客户转化率在10%左右，最好的一个月达到13%；C.交叉验证，使用每个月的实际数据作为验证数据集，提升图曲线画观察差距，差距并不大；
6. 信用卡至保险基金交叉销售模型：
7. 背景：花旗银行贵宾客户保险渗透率高达47%，基金渗透率高达56%，存量客户的开发已经趋于饱和，需要寻找新的增长点；信用卡客户和借记卡客户的交叉销售是为重点待发掘市场；
8. 模型建立：选取全体信用卡客户作为样本（46万），2017年11月之前的六个月作为观察窗口，2017年11月到2018年4月作为表现窗口，通过信用卡刷卡付保费，但未在我行购买保险的客户作为保险目标变量；同时持有借记卡和信用卡并购买了基金的客户作为基金目标变量；
9. 模型结论：选取决策树中的responder rate最高的7个树叶作为目标客群；客户的特点包括：
10. 保险：基本特征（人生阶段中处于发展阶段（年轻37岁，已婚，附属卡）， 社会阶层属于处于打拼、积累资产的白领人士（月收入中位数1万人民币，高等教育，有车但不一定有房）），行为模式（信用卡重度活跃用户，消费欲望旺盛，喜欢出国旅游购物，极高海外消费，海外消费均值和中位数是普通客户的10倍以上，消费习惯上喜欢逛商场，聚餐，贡献的ANR/CNR远高于普通用户）；

产品偏好（偏向于性价比较高的重疾险等健康保险）；特殊发现（风险等级为c的responder rate最高，原因估计是该类客户对未来有较大不确定、恐惧和担忧，因此有较长的保险保障动力）

1. 基金：基本特征(（人生阶段处于成熟阶段（年龄42岁，已婚，附属卡），社会阶层属于富裕阶层（月收入中位数2.2万人民币，高教育，有房族无房贷，职务为高管，企业主等，持有高端信用卡（Preimier, prestige, ultima））; 行为模式上同样是海外消费较高，经常出国，换汇行为较多）；
2. 模型表现：训练数据集的KS值分别达到51和36，提升图最大提升倍数为32倍、18倍（保险random rate 0.4%，最高responder rate 12%； 基金random rate 0.5%, responder rate 4%）；目前还在和业务部门讨论协调，报领导审阅，还没有执行；

注意：增加模型运行后转化率，测准率数据

1. 客户流失/存款流失模型（14年模型更新）
2. 模型背景：受6家分行关行、贵宾客户达标资产标准提高两个因素的影响，花旗银行2018年一季度50万以上客户流失非常严重，一季度流失客户2500户（净减少1500户，其中流失2500户，新获客加提升1000户），流失率10%，年化流失率将达到40%；为了在防止客户流失的同时推动客户资产提升到100万，更新流失预警模型，产出高危流失客户清单交一线重点跟进
3. 模型建立：选取全体借记卡客户作为样本（14万），以2018年1月之前的六个月作为观察窗口，2018年1月到3月作为表现窗口；2018年资产降低到50万以下的客户作为目标变量；通过数据准备、EDA数据分析、WOE粗分类等过程建立逻辑回归模型；
4. 模型结论：最终选取了5个变量，主要包括，产品持有（产品数量低于3个（不包含活期结算账户），投资产品AUM占比低于40%，未来三个月将要到期的产品占全部AUM的50%以上；行为特点（活跃度低，最近三个月没有发生主动交易（产品到期、利息到账不算主动交易），最近三个月未登录网银和手机银行；）
5. 模型表现：训练数据集KS值35，提升图最大提升倍数5倍数（random rate: 10%，responder rate 接近60%）；5月份投入运行，产出提升图前20%（约4400人，流失率高于50%），的人群下发一线营销挽留，推动客户资产提升；5月份流失客户数有所缓和，过去每个月流失830人，5月份流失600人）
6. 与2014年相比，主要变化指标：存款占比改为investment占比，产品max到期日30天以内改为未来到期资产占AUM比重；行龄（36个月以上客户易流失）改为客户活跃度；
7. MGM获客模型（14年模型更新）
8. 存款流失分析
9. 背景：花旗银行个人存款自2017年5月至18年5月下降10%，由21亿美元流失至18亿美元，同期中国个人存款总额增长7%。花旗银行在占有利率优势（利率同业最高）、市场优势（美元升值）的情况下存款却逆势下跌，内在原因需要查明。
10. 过程：

(1).和同业比，增势相反，明确是内部而非外部原因

(2).

1. 结论：在外部因素有利的签体下（市场、汇率）；花旗银行存款逆市下跌，主要由内部因素驱动：
2. 【网点调整】，关行，关行/及拟关行网点占流失存款的48%；
3. 【管户调整】另外52%的存款流失来自正常营业网点，这些网点80%以上的存款流失来自于PBRM change的客户，其中人员变化较大的北京下降最大，占全国的30%以上；
4. 【产品单一】超过70%流失存款来自单一存款客户；其中过半来自MOB小于1年的新客户；
5. 汇丰银行催收评分卡模型
6. 模型背景：汇丰银行目前按照逾期期数和金额人为地将其简单粗放地分为高中低风险三类，对不同群体采取不同的催收策略, 这种一刀切的分群方法过于依赖于管理人员的主观判断和经验，难以精准识别真正的高中低风险客户，导致催收资源错配、催收效率不佳：对低逾期期数/低金额的高风险客户催收不足，错过最佳催收时机使起滑入不良。对逾期金额高/但能够自行还款的优质客户又进行过度催收，做无用功的同时也打扰客户。汇丰需要一套基于历史数据的统计预测模型，精准识别高还款可能性客户，科学制定催收策略: 1.对自愈客户不予催收或延期催收（如逾期15天之后再催收）2.低风险客户分配给本行催收团队电话催收，高风险客户委外。3.根据客户风险程度（还款可能性）对催收员奖金、外包公司回收奖金实行阶梯定价，（提高高风险客户的佣金）激发队伍积极性。3.极高风险客户（还款可能性极低客户），除非金额巨大，否则不予催收（催了白催），节省人力。
7. 模型建立：选取全体信用卡逾期客户为样本（2019年3月），2019年3月之前的12个月为观察窗口，之后的3个月为表现窗口，在表现窗口中至少还最低还款额为目标变量，建立逻辑回归模型。最终5个应变量进入模型：逾期期数，逾期金额，过去6月信用额度用尽次数，历史还款行为……; 其中IV值最高的决定性变量是逾期期数。目前考虑增加更多外部数据变量，为什么？因为按照业务经验，客户是否还款取决于两点1.还款意愿2.还款能力，而还款能力很大程度取决于个人当时的经济状况、负债状况，人行征信等外部数据披露客户外部负债情况，(关键是DBR)，应该对是否还款有较高预测力。

理想的催收评分卡模型既能识别出坏客户，也就是还款可能性极低的客户，不予催收或选择性催收，也能识别出好客户，也就是还款可能性极高的客户，不予催收或延迟催收。

最终模型起到三个效果：1.优化催收人力配置，不做无用功；2.优化激励机制，激发队伍积极性。3.精细化催收策略，提高催收效果。(还款意愿+还款能力)

1. 模型表现：训练数据集、验证数据集KS均达到60。
2. 模型投入运行后成效：
3. 汇丰银行失联客户预测模型
4. 12341234
5. 1234234