**Wind落地数据库26家上市银行相关指标数据提取说明**

1. **bankdatacollection.mat:**
2. banklist（26家上市银行基本信息）：
3. 第一列：各银行对应代码；
4. 第二列：名称；
5. 第三列：上市时间；
6. 第四列：股票代码。
7. dateinfo（季报公告期）：

记录2005-2018年间季报公告时间（每年3月31日、6月30日、9月30日、12月31日），共53行数据。

1. factorinfo（收集的各个指标的相关信息）:
2. 第一列：各指标在datanew结构体中对应的名称；
3. 第二列：各指标名称；
4. 第三列：各指标在落地数据库中对应的编码；
5. data（包含各个指标所有数据信息的结构体）：

例如：data.ROE中，第一行的53个数据表示:banklist中第一行对应的银行，在dateinfo中的53个日期公布的53个净资产收益率数据。

以此类推，26行数据分别记录对应的26家银行53次季报中公布的数据。

1. getfactorinfo.m**：**

1.利用matlab调用SQL语言从落地数据库中获取相关指标的数据。

2.从获取的所有数据中挑选出26家银行的相关数据。

3.对数据整理、归类、命名。

**三、数据缺失情况**（以上市日期之后为准）：

1.缺失较少的指标**：**

**存放同业和其他金融机构款项、发放贷款及垫款可供出售金融资产、公允价值变动净收益、其他资产、同业和其它金融机构存放款项、卖出回购、其他负债、成本收入比、净息差、营业外利润、吸收存款**：已上市银行2005，2006年的数据部分缺失（大部分只有年报数据）。

**现金及存放中央银行款项、逾期贷款、平均净资产收益率：**大部分银行只有半年报数据。

**拆出资金**：已上市银行2005，2006年的数据部分缺失（大部分只有年报数据），宁波银行、无锡银行、贵阳银行上市后仍有部分数据缺失。

**交易性金融资产**：浦发银行缺2009和2010年数据，华夏、民生、招商缺2006年数据。

**可供出售金融资产**：已上市银行2005，2006年的数据部分缺失（大部分只有年报数据），18年第一季度部分银行数据缺失。

**长期投资股权**：平安、宁波、华夏、江苏、光大、贵阳后期数据全部缺失。

**应收款项类投资：**18年第一季度部分银行数据缺失；交通银行2007、2008年数据缺失，光大银行2010-2012年数据缺失（只有年报数据）。

**拆入资金**：已上市银行2005，2006年的数据部分缺失（大部分只有年报数据）；无锡银行17年报、吴江银行17年第二季度数据缺失。

**应付职工薪酬、买入返售金融资产**：已上市银行2005，2006年的数据部分缺失（大部分只有年报数据），中国银行07年数据全部缺失。

**应付债券、净利润同比增速**：已上市银行05-07年的数据大部分缺失。

**利息净收入、手续费及佣金净收入、资产减值损失**：已上市银行05年的数据大部分缺失。

**净资产收益率同比变动、总负债**：部分银行上市第一年的数据部分缺失（只有年报数据）。

**不良率、拨备覆盖率**：华夏银行05-11年的数据只有半年报数据，北京银行15-16年数据缺失，中国银行07年只有半年报数据，中信银行07-08年只有半年报数据，吴江银行17年只有半年报数据。

**资本充足率、核心资本充足率**：15-18年大部分银行数据几乎全部缺失。

**总存款：**华夏银行05-08年的数据只有半年报数据，大部分银行上市第一年的数据只有半年报。

2.缺失比较严重的指标**：**

**衍生金融资产、衍生金融负债**：7家银行没有相关数据，已上市银行2005，2006年的数据部分缺失（大部分只有年报数据）。

**向中央银行借款**：05-10年数据几乎全部缺失。

**贵金属、递延所得税负债**：一半的银行没有相关的数据。

**交易性金融负债、存款证、生息资产平均收益率：**大部分数据都缺失。