

Compustat on WRDS

Lili Liu
S&P Global Market Intelligence
Enterprise Solutions
Specialist

Nov 28 2018

清华大学 Compustat 培训

Lili Liu
S&P Global Market Intelligence
Enterprise Solutions
Specialist

Nov 28 2018

S&P Global

Market Intelligence

主要内容

- 关于Compustat
- Compustat 数据在沃顿平台
- Compustat 数据分类及覆盖（清华）
- Compustat 数据基础构成
- 客服服务



关于 Compustat

About us

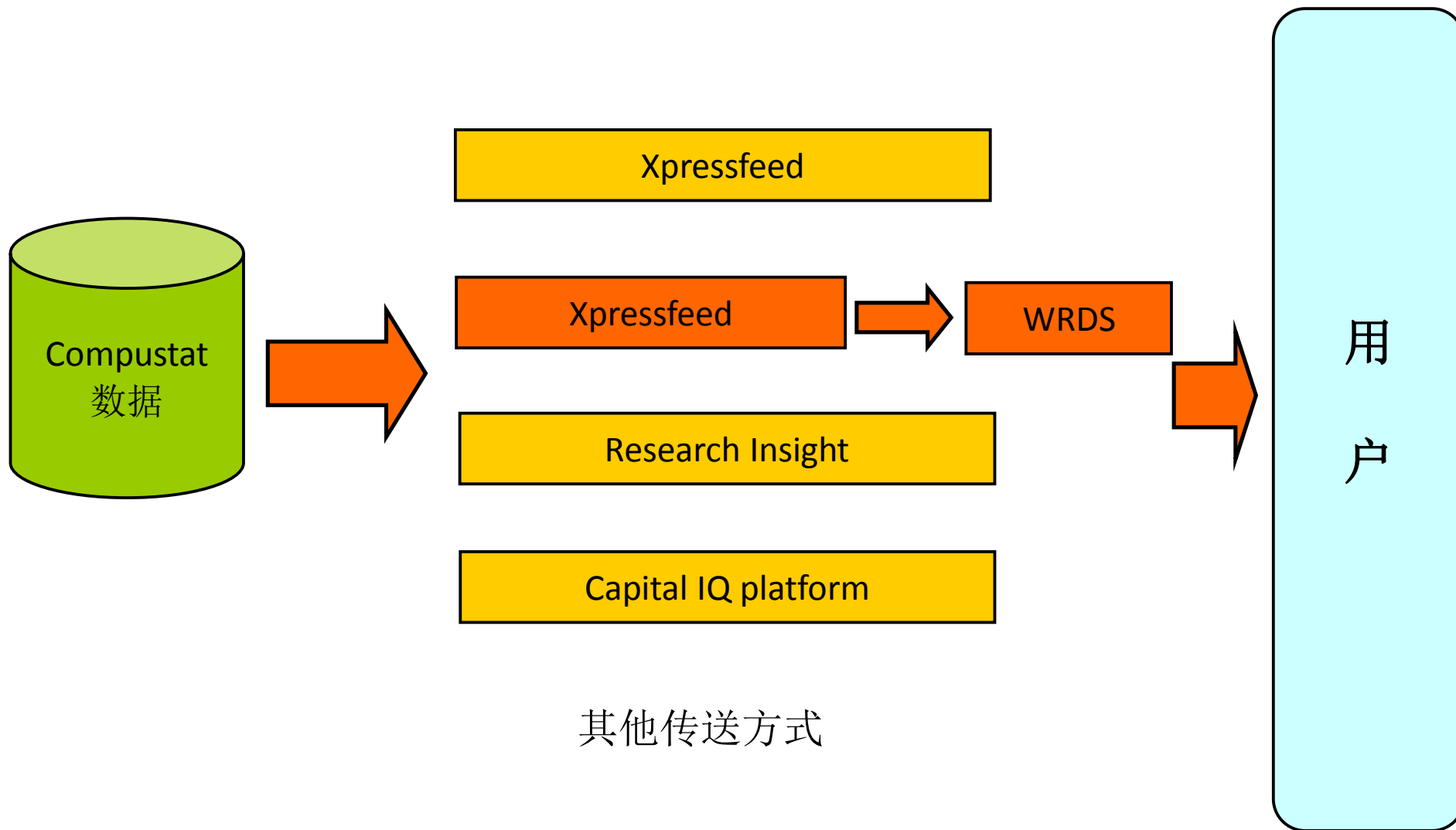
- **Compustat** 是一个全面的财务数据库, 收录公司的营运及财务状况资料库
- **Compustat** 涵盖130多个国家, 超过 104,000上市/已退市公司*
- 超过 130,000只证券, 占全球市值99%以上*
- 提供数据包括:
 - 财务报告, 市场价格, 指数价格
 - 行业分类代码 (如: **GICS**, **SIC**等)

*数据统计截至2018年07月

About us

- **Wharton Research Data Services (WRDS)** - 是我们的合作伙伴, 提供学校一个平台获取Compustat数据
- 现在, 全球有超过320家顶级大学经过 WRDS 订约了 Compustat, 包括:
 - Peking University
 - Tsinghua University
 - University of California, Berkeley
 - University of Hong Kong
 - The Hong Kong University of Science and Technology
 - Chinese University of Hong Kong
 - Boston University
 - University of Sydney
 - Massachusetts Institute of Technology...

关于Compustat





Compustat 数据分类及覆盖

Compustat 数据分类及覆盖

- North America & Backdata
- Global

North America & Backdata

Fundamentals Annual

Fundamentals Quarterly

Index Constituents

Index Fundamentals

Index Prices

Industry Specific Annual

Industry Specific Quarterly

Pension Annual

Pension Quarterly

Ratings

Security Daily

Security Monthly

- **Compustat North America & Backdata** 提供美国和加拿大上市公司及已退市公司的基础财务数据，市场交易数据，公司养老金数据，标普指数和标普评级信息
- 数据覆盖 - 46,000多家公司*
- 数据更新频率 - 每月更新 (monthly)
- 数据历史 -
 - 年度数据可以回溯到1979年，其中有35个科目年度历史数据可回溯到1950年
 - 季度数据提供12年历史，其中35个科目可追溯到1962年
 - 月度交易价格追溯到1962年
- 数据科目
 - 超过300个年度数据科目和100个季度数据科目，覆盖利润表、资产负债表、现金流量表及附注科目
- 数据格式
 - 年度：历史格式和经调整的格式 (Restated)
 - 季度：经调整的格式

*数据统计截至2018年07月

North America & Backdata

Fundamentals Annual

Fundamentals Quarterly

Index Constituents

Index Fundamentals

Index Prices

Industry Specific Annual

Industry Specific Quarterly

Pension Annual

Pension Quarterly

Ratings

Security Daily

Security Monthly

- 特殊行业包含

- Airlines

- Gaming

- Healthcare Facilities

- Home Building

- Lodging

- Managed Care

- Mining

- Oil & Gas

- Oil & Gas Refining

- Retail

- Semi-conductors

- Bank & Thrift

North America & Backdata

Pension Annual

Pension Quarterly

Ratings

Security Daily

Security Monthly

Segments (Non-Historical)

Segments (Non-Historical)

- Customer

- 养老金：主要披露公司为员工退休后提供的福利，分为退休前的养老金缴存计划和退休后公司提供的福利。
- 缴存计划：由于美国的会计准则变化，数据可追溯的历史不同，但基本都可以追溯到1987左右。
- 退休员工福利：由于美国的会计准则变化，数据可追溯到1993年左右。

North America & Backdata

Pension Annual

Pension Quarterly

Ratings

Security Daily

Security Monthly

Segments (Non-Historical)

Segments (Non-Historical)
- Customer

企业细分数据： 提供美国和加拿大公司根据不同的业务和区域细分数据

- 数据为年度数据
- 覆盖面—超过 31,000 家公司
- 数据历史—可回溯到1978
- 数据来源—企业披露的年报

North America & Backdata

- 举例

S&P Global 在2017年年报中披露的各业务部门的销售收入及成本:

	gvkey	conml	stype	SegmentName	datadate	sales	cogss
1	007163	S&P Global Inc	BUSSEG	Ratings	2017-12-31 00:00:00.000	2878	850
2	007163	S&P Global Inc	BUSSEG	Indices	2017-12-31 00:00:00.000	733	139
3	007163	S&P Global Inc	BUSSEG	Market & Commodities Intelligence	2017-12-31 00:00:00.000	2452	833
4	007163	S&P Global Inc	BUSSEG	Corporate	2017-12-31 00:00:00.000	0	-109

Global data

Compustat - Capital IQ

Global - Daily

Fundamentals Annual

Fundamentals Quarterly

Index Constituents

Index Prices

Security Daily

- **Compustat**全球数据收录北美以外的其他区域公司，包含基础财务数据及市场交易数据
- 数据调整: **Compustat Global**数据是经过标准化设定的，以便在不同国家和不同会计准则下面做比较
- 数据覆盖: 超过**6**万家公司，占全世界市值的**90%**以上*
- 历史:
 - 年度基础财务数据可追溯到**1979**年
 - 季度基础财务数据可追溯到**2003**年第四季度
 - 每天交易价格可追溯到**1989**年

*数据统计截至2018年07月

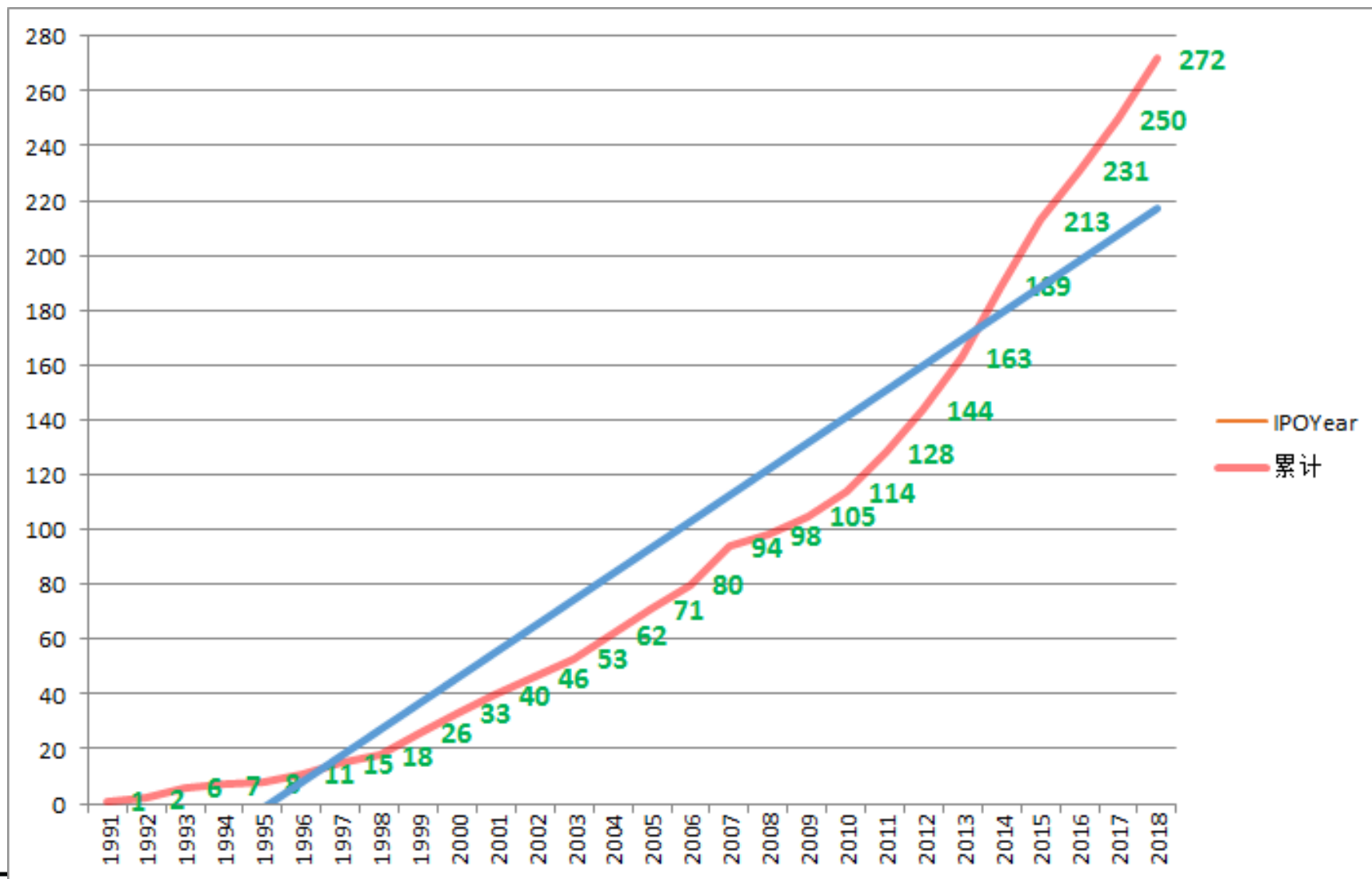
S&P Global

Market Intelligence

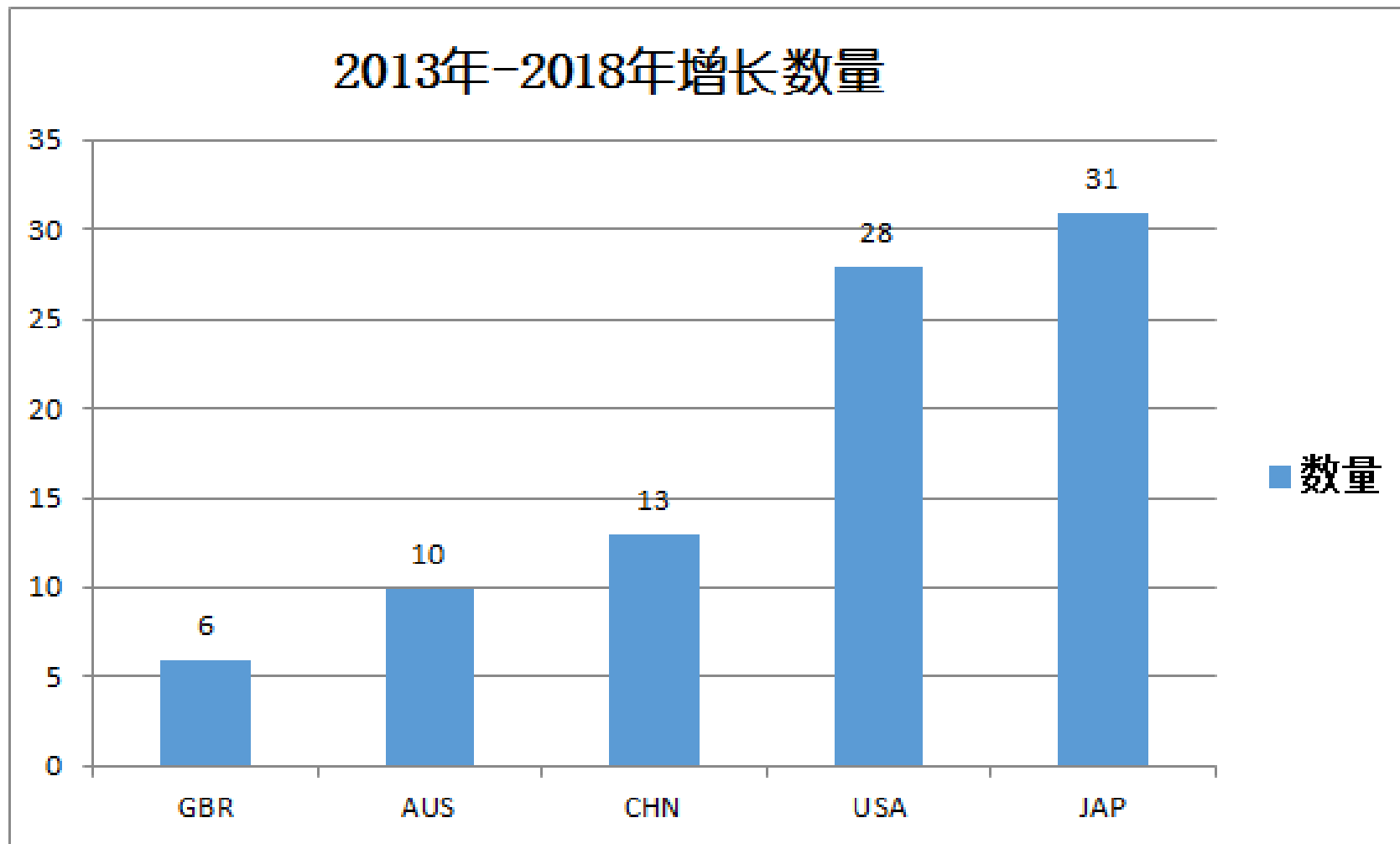
数据库的应用

- 投行的一般应用：
 - 行业的分析：以GICS子行业分类的Interactive Media & Services（50203010）这些高科技企业在今年发展迅速，且中国公司的比例在慢慢加大。
 - 公司的盈利情况，基本面的信息，预测公司的发展状况
- 基金公司的应用：量化分析，主被动投资。Compustat和CIQ的数据也可用于量化投资，根据对市场的研究，制定自己的量化因子，通过数据库做回测，并在回测中调整自己的投资策略。
- 四大会计师事务所：行业分析，并购策略，Compustat和CIQ数据库也关注与国际税收，转让定价这部分的数据。

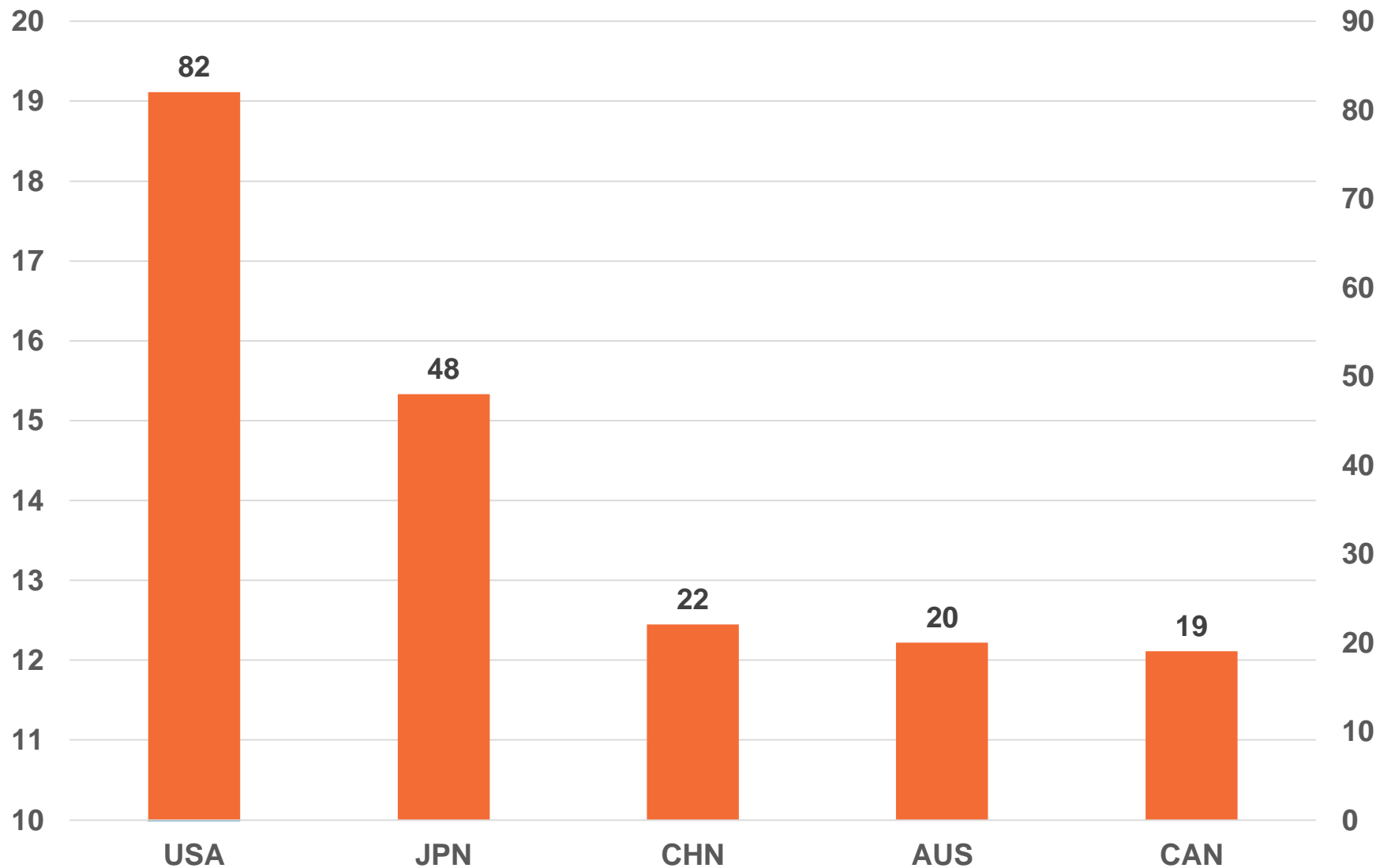
50203010行业的上市公司数量



2013-2018年主要国家增长的公司数目



前5个国家公司总数





Compustat 数据基础构成

Compustat 数据基础构成

- **Compustat 数据：**
 - 公司层面数据 (gvkey)
 - 季度财务数据
 - 年度财务数据
 - 股票层面数据(gvkey和iid)
 - 股票的价格数据
 - 股本及交易量
 - 拆分及分红信息
 - 指数层面数据(gvkeyx)
 - 指数成分股
 - 指数价格

公司层面

- 基本信息

- 唯一识别码(Gvkey)
- 公司名称(CONM)
- 业务描述(Busdesc)
- 是否还是活跃公司(Costat)

- 财务数据描述信息

- 数据更新代码 (UPD)
 - ✓ 2表示当前财务数据是初步财务数据的披露
 - ✓ 3表示最终的财务数据披露
- 数据来源(SRC)
- 财务年度截止月份(FYR): 5月31日为分割日期
- 财务报表披露使用的货币(CURCD), 数据的单位是以百万计
- 披露时所用的会计准则(ACCTSTD)
- 会计准则变化(ACCTCHG)

公司层面

• 6个主要制定报告的科目

- **GVKEY**: 公司唯一识别码
- **Datadate**: 数据时间
- **Consolidation**:
 - 描述报告是母公司数据还是与子公司合并的数据
 - **C**: 表示合并报告，摘录于公司报告
 - **N**: 表示母公司报告
- **Data Format**:
 - 表示数据收集及呈现的格式
 - 北美公司: **STD, SUMM_STD**
 - 全球公司: **HIST_STD, RST_STD**
- **Industrial Format**:
 - 表示公司是金融类(**FS**)或者工业类(**INDL**)
- **Population Source**:
 - 表示样本的来源 本土(**D**) 还是国外 (**I**)
 - **D**: 表示美国、加拿大和 **ADR**公司
 - **I**: 表示北美以外的公司

公司层面财务数据

举例：

JPMorgan 在2017财年资产总额摘录的3个记录

第一个是金融类的标准格式

第二个是工业类的标准格式

第三个是工业类的经调整过的格式

	CompanyName	gvkey	datadate	indfmt	datafmt	consol	popsrc	Asset_Total
1	JPMORGAN CHASE & CO	002968	2017-12-31 00:00:00.000	FS	STD	C	D	2533600
2	JPMORGAN CHASE & CO	002968	2017-12-31 00:00:00.000	INDL	STD	C	D	2533600
3	JPMORGAN CHASE & CO	002968	2017-12-31 00:00:00.000	INDL	SUMM_STD	C	D	2556062

脚注（Footnote）

Compustat 数据的一个重要组成部分

用字母来表示财务数据的收集情况

常用脚注解释：

- AC 由于会计准则变化，财务数据做出了调整
- AA 当年数据受到并购影响
- **BD 销货成本COGS调整了折旧和摊销（不包含）**
- GB 资产负债表中的科目没有细分
- IC 数据科目中包含从长期负债披露中获得的流动负债

以苹果2017年年度数据为例

数据摘录于年报中的主表，附注1和4

	gvkey	conml	datadate	cogs	cogs_fn
1	001690	Apple Inc	2017-09-30 00:00:00.000	131648	BD

2017

Net sales \$ 229,234

Cost of sales 141,048

Property, Plant and Equipment

Property, plant and equipment are stated at cost. Depreciation is computed by use of the straight-line method over the estimated useful lives of the assets, which for buildings is the lesser of 30 years or the remaining life of the underlying building; between one and five years for machinery and equipment, including product tooling and manufacturing process equipment; and the shorter of lease term or useful life for leasehold improvements. The Company capitalizes eligible costs to acquire or develop internal-use software that are incurred subsequent to the preliminary project stage. Capitalized costs related to internal-use software are amortized using the straight-line method over the estimated useful lives of the assets, which range from three to five years. Depreciation and amortization expense on property and equipment was \$8.2 billion, \$8.3 billion and \$9.2 billion during 2017, 2016 and 2015, respectively.

Amortization expense related to acquired intangible assets was \$1.2 billion, \$1.5 billion and \$1.3 billion in 2017, 2016 and 2015, respectively. A amortization expense related to acquired intangible assets as of September 30, 2017, is as follows (in millions):

2018

- $COGS - DP = 141048 - 9400 = 131648$
- $DP = 8200 + 1200 = 9400$

数据代码（Datacode）

数据代码是指出为什么个别科目没有数据。

数据代码主要有以下表示方法：

- 3- AF = Annual Figure（年度数据）
- 4- CF = Combined Figure（合并数据）
- 8- IF = Insignificant Figure（非重要数据）
- 2- SF = Semi-Annual Figure（半年度数据）

股票层面

- 基本信息
 - 识别码: gvkey+iid, CUSIP, ISIN, SEDOL*, Tic
 - 是否还是活跃股票(secstat)
 - 不活跃的时间(dldtei)
 - 不活跃的原因(dlrsni)
 - 交易所信息(exchg)
- 价格数据
 - 高开低收 (prchd,prcod,prcld,prccd)
 - 股本(cshoc)
 - 交易量(cshtrd)
 - 拆分因子(ajexdi)
 - 没有价格的原因(prcstd)

指数层面

基本信息

- 识别码: gvkeyx
- 指数名称(conm)
- 所在区域(indexgeo)
- 归属于那个指数公司(indexcat)
- 是否包含成分股(idxcstflg)
- 是否活跃(idxstat)
- 指数价格(prccd)



客户服务

客户服务

- 客户服务

邮箱: support.datafeed.mi@spglobal.com

电话: 8610 6569 2985

中文客服周一至周五, 早8点到晚5点

英文客服24小时7天

谢谢！