

大数据分析 IT 行业薪资变化及影响因素

顾滢双

大连理工大学，辽宁省 大连市 116000

摘要：为了帮助互联网人才更好地了解行业动态趋势，文章利用大数据技术对 IT 行业薪资变化及影响因素进行分析。文章基于从各大线上招聘平台爬取的数据以及从各大开源数据库中获取的数据，利用相关性分析、可视化分析等方法，从多个角度分析薪资影响因素。文章最后对 IT 行业薪资变化趋势进行了总结，以希为广大 IT 从业者提供行业前景展望，同时为所有想进入 IT 行业的求职者提供参考。

关键词：人才就业；互联网薪资；变化趋势；影响因素

1. 引言

作为国家与企业发展的根本，人才问题一直是社会发展关注的热点问题。随着互联网技术的不断发展，“互联网+”时代已然到来，各大企业对互联网相关人才的人才需求缺口更是逐年扩大。研究显示，2017 年中美两国的互联网相关人才需求量达到了供应量的 10 倍之多^[1]。

在当今社会，如何招聘到满意的人才填补人才缺口，已经成为各大企业人力资源部门要重点解决的问题。而薪资作为广大求职者最关心的问题，在解决企业人才招聘、行业就业问题中，都发挥了举足轻重的作用^[2-3]。本文的数据来源于国内某线上招聘网站上 2019 年 6 月份发布的上万条 IT 行业招聘数据，利用爬虫技术获取这些数据，从多个维度对行业薪资的影响因素进行相关性、可视化分析，希望能为 IT 求职者和各大企业的 IT 人才招聘提供帮助与参考。

2. 分析结果

2.1 IT 行业薪资变化

首先对 IT 行业的薪资变化情况做总体分析，图 1 以 IT 行业的代表性职位，前端开发工程师的平均月薪（k/月——千/月，下文同义）代表 IT 行业的平均薪资，显示其自 2014 年到 2021 年的薪资变化情况。可以看到，IT 行业的薪资在 2014 年到 2018 年每年呈显著上升趋势，从 2018 年到 2020 年呈缓慢增长趋势，这与 IT 行业的不断发展，人才缺口的不断扩大的社会趋势相符合。从 2020 年到 2021 年的平均薪资有所下降，这与全球新冠肺炎疫情有很大关系，由于疫情，许多企业对 IT 人才的招聘需求有所放缓，因此行业平均薪资有所下降，属于正常行业波动。

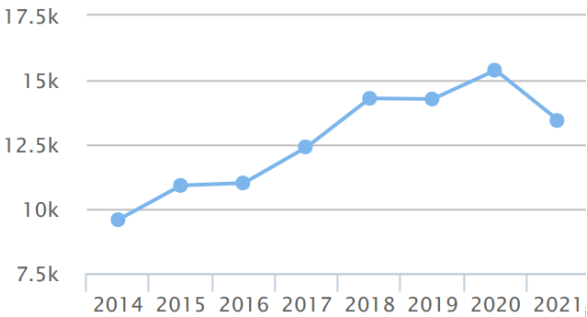


图 1：IT 行业平均薪资变化趋势

2.2 IT 行业薪资分布总体情况

图 2 显示了 IT 行业薪资的总体分布情况，数据显示，各企业提供的 IT 岗位的平均薪资约为 17.5k/月，约有 11000 个岗位薪资超过了平均薪资，约占总职位数的 40%。由图 2 可见，IT 行业的薪资分布总体位于 10k/月到 23k/月的区间内，但也有少量高薪职位薪资可达到 30k/月及以上。

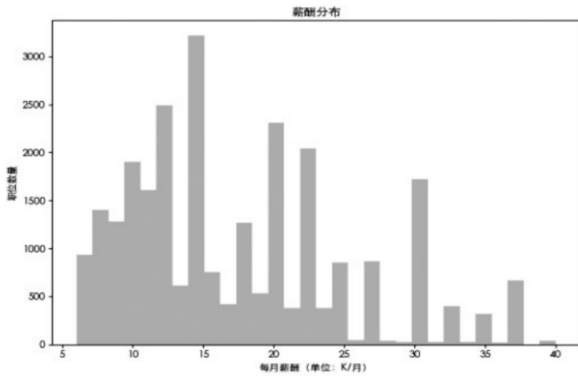


图 2 薪资水平分布总体情况（月薪：k/月）

2.3 IT 行业薪资与企业所在城市的关系

不同城市对 IT 行业人才的需求量不同，因此提供的 IT 岗位平均薪资也有显著不同。图 3 显示了对 IT 人才需求量排名前 10 的城市以及这些城市的 IT 行业薪资情况。可以看到薪资较高的基本集中在一线城市，如北上广深，和新一线城市，如杭州成都等。由图 3 可见，北京的 IT 行业薪资最

高，中位数为 20k/月，杭州的平均薪资位居新一线城市首位，中位数接近 18k/月。有这样趋势的原因，一方面，一线及新一线城市 IT 行业发达，IT 企业数量多，对 IT 人才的需求量大，并且大多有国家的政策扶持。北上广深集中了许多 IT 行业龙头企业的总部，有良好的 IT 就业氛围。杭州今年来在阿里巴巴的带动下，IT 产业迅速发展，因此薪资水平较高。

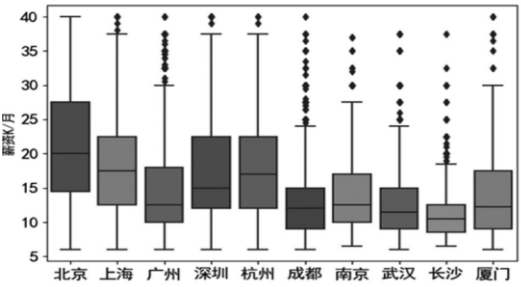


图3 IT行业薪资水平前10的城市

2.4 IT行业薪资与企业规模的关系

企业规模能够反映企业的财政情况，同时也会影响招聘机制，因此不同规模的企业提供的薪资水平也不同。图4反映了IT行业薪资与企业规模的关系，可以看到随着企业规模的扩大，薪资水平总体呈上升趋势。规模在150人以下的企业属于中小型企业，提供的平均薪资相对较低，但薪资上限有可能超过大规模企业的最大值，这也体现了中小型企业虽然财力不如大规模企业，但用人机制相对灵活，薪资幅度较大。规模在500人以上的大规模企业平均薪资能达到20k/月，但薪资幅度相对稳定，这也符合大规模企业体制完善的特点。

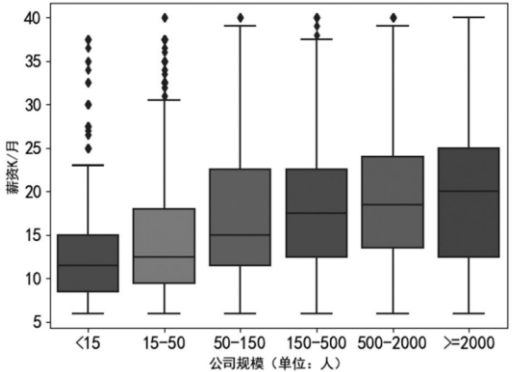


图4 IT行业薪资与企业规模的关系

2.5 IT行业薪资与企业融资情况的关系

图5反映了企业融资规模对IT行业薪资的影响。未融资和天使轮的企业属于种子企业，企业财力

较弱，且用人风险较高，平均薪资约为12k/月；处于A轮、B轮、C轮融资后的企业逐渐步入正轨，愿意提供较高的薪资水平吸引人才推动企业发展；D轮及以上的公司薪资最高，达到23k/月。然而，上市公司的薪资低于C轮或C轮以上的公司，可能是因为上市企业整体发展趋于平稳，工作强度适中，工作较为稳定，求职者愿意为此放弃更高的薪资。

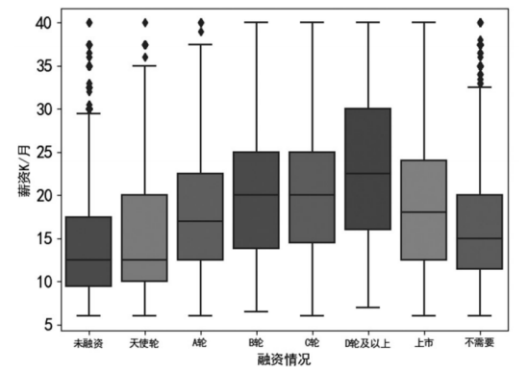


图5 企业融资规模对IT行业薪资的影响

2.6 IT行业薪资与岗位类型的关系

图6反映了IT行业薪资与岗位类型的关系，数据分析岗的薪资水平最高，在15-25k/月。这与近年来各大企业意识到数据挖掘、分析在市场上的重要作用，加大对数据分析人才的投入有关。作为传统的互联网岗位，研发和产品依然是求职者投递的热门，其薪资也相对较高；设计与运营岗的薪资幅度较大，上限甚至可能超过研发、产品和数据分析，但依赖于具体的方向，例如平面设计师和新媒体运营岗位的薪资会较高。

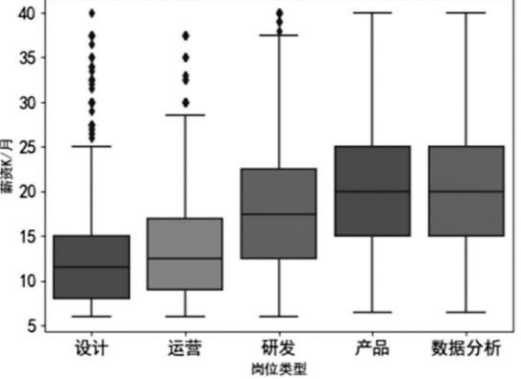


图6 IT行业各岗位的平均薪资

2.7 IT行业薪资与工作经验的关系

图7反映了IT行业薪资与工作经验的关系，总体来看，随着工作经验的增加，薪资水平呈上升趋势。3年以下的工作经验薪资水平没有明显差异，

这是因为在 IT 行业，3 年工作经验才能基本入门，对所在行业和领域有成熟的认识。具有 10 年以上工作经验的，薪资中位数可达 30k/月，企业对此类经验丰富的人才需求较大，因此薪资较高。

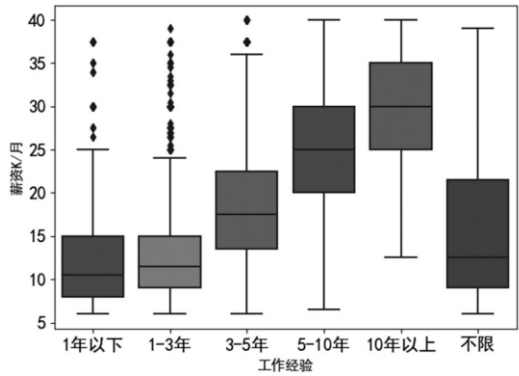


图 7 工作经验对 IT 行业薪资的影响

2.8 IT 行业薪资与学历的关系

图 8 反映了 IT 行业薪资与学历的关系，可见在 IT 行业学历越高薪资水平明显越高，本科生的薪资相比于大专生，涨幅将近 50%，博士生的薪资中位数更是达到了 30k/月。

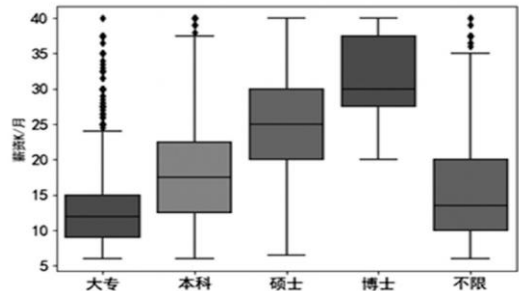


图 8 学历对 IT 行业薪资的影响

2.9 IT 行业的基本岗位要求

图 9 的词云反映了企业对 IT 人才的关键要求，由图可见，团队合作、沟通、经验、技术、设计、应用等能力是企业最为看重的。



图 9 IT 行业岗位关键技能要求词云

3. 分析方法

首先利用 python 的 matplotlib、jieba 等可视化工具库，对爬取的数据进行可视化分析，生成上述图表，从多个维度对薪资影响因素进行具体直观的分析。

然后对薪资的影响因素进行具体的量化分析，计算薪资与各影响因素间的 Pearson 相关系数^[4]，这是一种用于度量两个变量间关联程度的方法，在风险预测^[5]、舆论评估^[6]等领域得到广泛应用。由于本文数据共 3 万多条，每条数据由求职薪资、公司名称、所在地、公司规模、融资情况、招聘岗位、经验要求、学历要求等 9 个变量组成，因此将求职薪资作为因变量，其余 8 个特征作为自变量，这 8 个特征的详细信息见表 1。

表 1 变量说明

变量类型	变量名	数据类型
企业基本信息	企业名称	文本数据，如：优信集团
		团
企业基本信息	企业所在地	定性变量，如：北京
企业基本信息	企业规模	定性变量，如：50-150人
		人
企业基本信息	融资情况	定性变量，如：A 轮
岗位需求信息	岗位名称	文本数据，如：Java 开发工程师
		发工程师
岗位需求信息	经验要求	定性变量，如：1-3 年
岗位需求信息	学历要求	定性变量，如：本科
岗位需求信息	岗位要求	文本数据，如：团队合作

Pearson 系数计算公式如下：

$$r = \frac{N \sum x_i y_i - \sum x_i \sum y_i}{\sqrt{N \sum x_i^2 - \left(\sum x_i\right)^2} \sqrt{N \sum y_i^2 - \left(\sum y_i\right)^2}}$$

r 的绝对值越接近 1，两个因素的相关性越强。利用 python 计算 9 个变量两两间的 Pearson 系数，并绘制热力图模型^[7]。从图 10 中可看出对薪资水平影响的最大的因素是工作经验，然后依次是城市、学历、岗位类型、融资情况和公司规模。

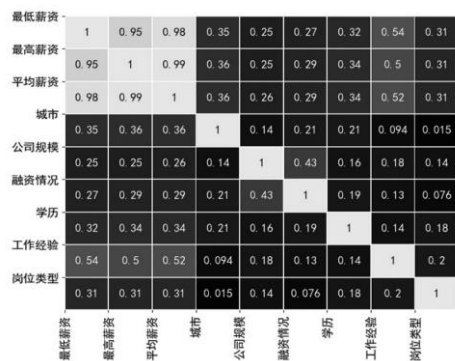


图 10 薪资影响因素关联性热力图

4. 结论

本文通过 python 爬虫技术，从国内某线上招聘平台爬取了 3 万多条招聘数据，通过数据可视化分析，获得了以下结论：①IT 行业薪资总体呈逐年上升趋势，近两年上升趋势放缓，且受疫情影响出现小幅下降波动；② IT 行业的薪资分布总体位于 10k/月到 23k/月的区间内，相较于其他行业有明显优势；③IT 行业薪资最高的城市集中在一线和新一线城市，有大型 IT 企业聚集的城市更容易形成良好的 IT 就业环境；④中小型企业平均薪资低于 500 人以上大规模企业，但其用人机制更灵活，薪资上限更高；⑤融资规模越大薪资水平

越高，但上市企业薪资趋于稳定；⑥数据分析岗位逐渐成为热门，薪资水平逐年升高，传统研发技术岗位薪资竞争力不减；⑦工作经验和学历越高，薪资水平有明显提升；⑧IT 企业十分注重求职者的团队协作沟通等能力。最后，本文计算了薪资水平和各影响因素间的 Pearson 系数，并据此绘制热力图，得出影响薪资水平的主要因素依次是：工作经验，企业城市，学历，岗位类型，企业融资情况。今后的研究方向，是结合机器学习理论，构建并训练模型，为每种类型的求职者预测其可能的薪资水平，为广大 IT 求职者提供求职建议和能力提升方向；并预测结合当今经济政治环境，对行业未来薪资变化进行预测，以求为广大 IT 人才提供行业风向参考。

5. 参考文献

- [1] 赵齐加. 全球互联网人才发展概况[J]. 今日科苑, 2018. 4:31-39
- [2] 秦浩, 郭薇. 国外人力资源服务业的发展现状及趋势[J]. 商业时代, 2013. 8:122-123
- [3] 陈远瞩. 商业银行应聘人员期望薪资的影响因素[D]. 华东师范大学, 2018.
- [4] 陈希孺. 概率论与数理统计[M]. 合肥: 中国科学技术大学出版社, 2009.
- [5] 赵海龙, 张丹丹, 黄松, 莫石, 魏浩. 基于皮尔逊相关系数的海南省地闪密度与雷击故障关系分析[J]. 高压电器, 2019. 8:186-192
- [6] 覃玉冰, 邓春林, 杨柳. 基于皮尔逊相关系数的网络舆情评估指标体系构建研究[J]. 情报线索, 2018. 10:15-19
- [7] 邱南森. 数据之美[M]. 中国人民大学出版社, 2014.