

2021

中国大数据产业发展指数 白皮书

北京大数据研究院

大数据分析与应用技术国家工程实验室

2021年10月

目录

一、 研究背景与指标体系	1
(一) 研究背景	2
(二) 指标体系	3
(三) 数据来源	4
(四) 计算过程	5
二、 大数据产业发展指数总体评价	6
(一) 大数据产业发展前二十强城市	7
(二) 大数据产业发展领先城市对比	9
(三) 大数据产业发展区域对比	10
三、 大数据产业发展分指数评析	12
(一) 产业政策与环境	14
(二) 产业规模与质量	16
(三) 头部企业情况	19
(四) 产业创新能力	22
(五) 产业投资热度	24
四、 展望与建议	26

一、研究背景与指标体系



一、研究背景与指标体系

(一) 研究背景

“十三五”期间，我国大数据产业迅猛发展，政策环境日益优化，产业规模稳步提升，产业价值充分释放，管理机制不断健全。“十四五”规划将“加快数字化发展，建设数字中国”作为独立篇章，明确指出大数据是七大数字经济重点产业之一，这是国家自《促进大数据发展行动纲要》颁布以来，对大数据发展做出的又一重要战略部署，为以大数据为重点的数字产业带来了新的发展契机。与此同时，自2020年3月30日中央指出加快培育数据要素市场以来，让数据成为新的关键生产要素，为产业赋能已成为发展趋势。数字助力提升传统产业技术创新能力，逐步打造创新驱动型的新兴大数据产业集群，正在成为稳步推进大数据产业向高端化、智能化方向迈进的重要趋势。

为紧跟国家政策方针，明晰各城市大数据产业的发展短板及存在差距，团队在2020年发布大数据产业发展指数的工作基础上，进一步扩充了调研企业数量，并跟进调研各地大数据产业、大数据企业发展状况和政策环境，继续编制了2021年大数据产业发展指数，力求为各城市大数据产业发展提供方向性指引，发挥数字经济的新引擎作用，为数字中国建设赋能添力。

一、研究背景与指标体系

(二) 指标体系

依托2020年大数据产业指数的相关研究成果，本年度仍从产业政策与环境、产业规模与质量、头部企业情况、产业创新能力、产业投资热度五个维度展开评价，共涉及13个二级指标与22个三级指标。

大数据产业发展指数指标体系

一级指标	二级指标	三级指标
产业政策与环境	政策环境	大数据政策发布情况
	组织建设	大数据管理机构设置情况
	战略布局	国家大数据综合试验区布局情况
	支撑服务	大数据产业联盟、协会及研究机构建设情况
产业规模与质量	产业规模	本地区大数据企业数量
	企业质量	总注册资本
		大数据企业网站建设情况
头部企业情况	上市企业	人员总规模
		上市企业数量
		上市企业总市值
		上市企业总净利润
	独角兽企业	上市企业平均技术人员占比
		独角兽企业数量
	瞪羚企业	独角兽企业总估值
		瞪羚企业数量
高新技术企业	高新技术企业数量	
产业创新能力	研发投入	企业R&D经费投入情况
	知识产权	企业软件著作权数量
		企业专利数量
		企业商标数量
产业投资热度	融资情况	企业总融资额
		企业总融资轮数

一、研究背景与指标体系

(三) 数据来源

2021大数据产业发展指数数据来源于北京大数据研究院企业库、北大法宝政策数据库以及相关政府信息公开信息。

1. 大数据企业数据库

北京大数据研究院大数据企业库入库企业为具备一定影响力或发展潜力的，从事数据生产、采集、存储、加工、分析、服务相关业务中的一环或多环的相关企业，截至目前共收录全国大数据企业6691家。企业库涵盖了122个维度指标，包括企业工商注册、运营情况、研发情况、投融资情况、产品情况等，并对大数据产业中的上市公司、独角兽企业和瞪羚企业，建立了头部企业库和产品库。基本情况如下：



2. 大数据政策库

团队与北大法宝合作，收录了全国3164项大数据相关政策，包括政策名称、发布部门、发文文号、发布日期、实施日期、时效性（现行有效、失效、已被修改等）、效力级别（行政许可批复、地方性法规、地方政府规章、地方规范性文件、地方工作文件等）等。

一、研究背景与指标体系

(四) 计算过程

1. 数据处理

用极差正规化法对数据进行无量纲化处理。

$$\text{正向指标标准化: } x'_i = \frac{x_i - \min x_i}{\max x_i - \min x_i};$$

$$\text{负向指标标准化: } x'_i = \frac{\max x_i - x_i}{\max x_i - \min x_i}。$$

2. 确定权重

利用“改进向量法”对指标体系中各指标赋予权重。即通过比较各样本特征向量到最优向量的距离，越小说明样本越接近最优水平，排名越靠前。距离采用欧式距离表示，如：在含三个特征（x,y,z）情况下，点A和B之间的欧式距离为：

$$\text{点AB间欧氏距离} = \sqrt{(x_a - x_b)^2 + (y_a - y_b)^2 + (z_a - z_b)^2}$$

3. 计算结果

基于得到的权重计算各城市的得分和排名情况。

$index_k$ 代表大数据产业指标体系中某级某个指标得分

$$index_k = \sum_{i=1}^{n_k} w_{ki} * x'_{ki}。$$

二、 大数据产业发展指数总体评价



二、大数据产业发展指数总体评价

（一）大数据产业发展前二十强城市

总体来看，大数据产业发展指数排名靠前的城市以直辖市、省会城市以及一线（新一线）城市为主，这些城市的大数据产业发展环境良好，产业发展势头良好。

在前10强城市中，长三角占5席、京津冀占2席、珠三角占2席、成渝经济圈占1席，东北和西北地区没有城市入围，表明大数据产业在全国各区域间的发展差异明显。从20强城市得分来看，长三角地区城市较为均衡，成都、重庆、贵阳引领西部地区大数据产业发展，京津两城的大数据产业发展差距与上一年度相比进一步拉大。

大数据产业发展指数排名（前二十名）

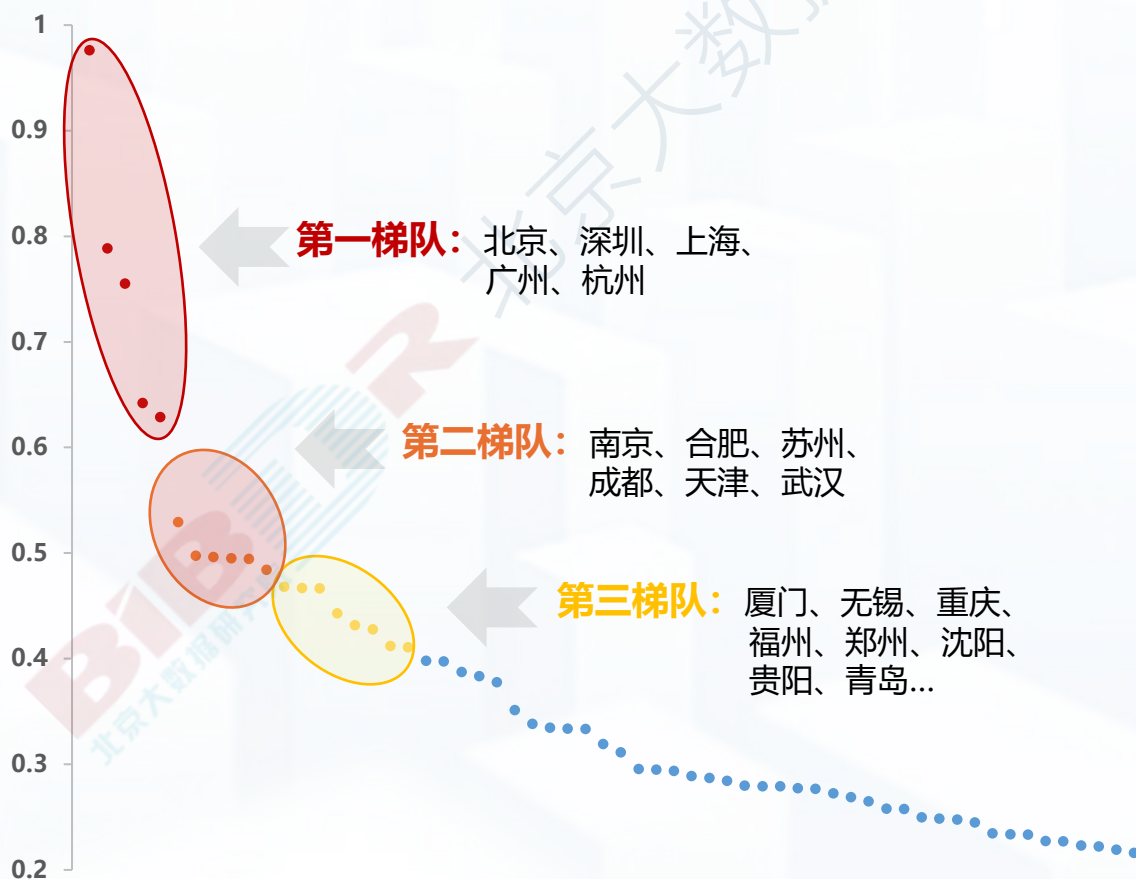
城市	2021 指数得分	2021 指数排名	2020 指数排名	城市	2021 指数得分	2021 指数排名	2020 指数排名
北京市	0.97585	1	1	武汉市	0.48378	11	7
深圳市	0.78815	2	2	厦门市	0.46793	12	13
上海市	0.75499	3	3	无锡市	0.46667	13	17
广州市	0.64193	4	4	重庆市	0.46632	14	11
杭州市	0.62840	5	5	福州市	0.44258	15	16
南京市	0.52902	6	6	郑州市	0.43140	16	15
合肥市	0.49716	7	12	沈阳市	0.42739	17	22
苏州市	0.49611	8	10	贵阳市	0.41159	18	14
成都市	0.49473	9	9	青岛市	0.41037	19	19
天津市	0.49401	10	8	长沙市	0.39792	20	20

二、大数据产业发展指数总体评价

(一) 大数据产业发展前20强城市

从梯队和排名变化情况可以看出，领先城市的排名地位稳固，由北京、深圳、上海、广州、杭州组成的第一梯队排名没有发生变化。与上一年度相比，由合肥、苏州、成都、重庆、武汉等城市组成的第二梯队内部竞争相对激烈，排名位次变化较大，其中武汉、重庆、贵阳排名下滑较大，合肥、无锡、沈阳等城市排名有所上升，合肥、苏州在政策引导、产业创新等多重因素的助推下，大数据产业发展整体趋势较好。

指数总排名散点图



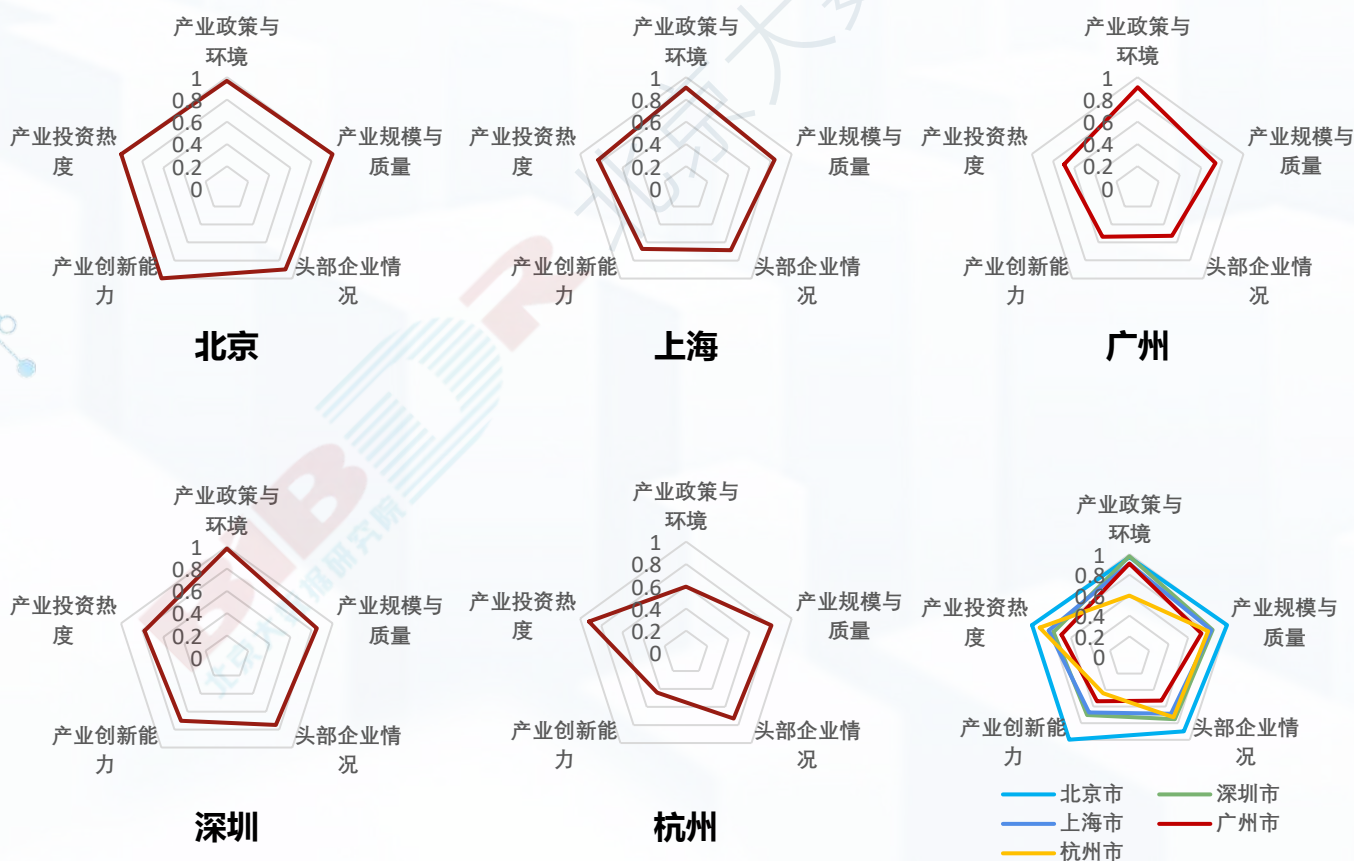
二、大数据产业发展指数总体评价

(二) 大数据产业发展领先城市对比

一线城市发展均衡，杭州大数据产业规模较小但投资热度高。

横向对比领先城市可以发现，北京在各项分指标上都表现突出，北上广深4个一线城市在各项指标的得分相对均衡。需要指出的是，杭州在投资热度方面格外突出，但在产业创新能力和产业政策与环境方面存在短板，在政策制定、机构设置、战略布局等方面仍然与一线城市存在差距。

领先城市对比



二、 大数据产业发展指数总体评价

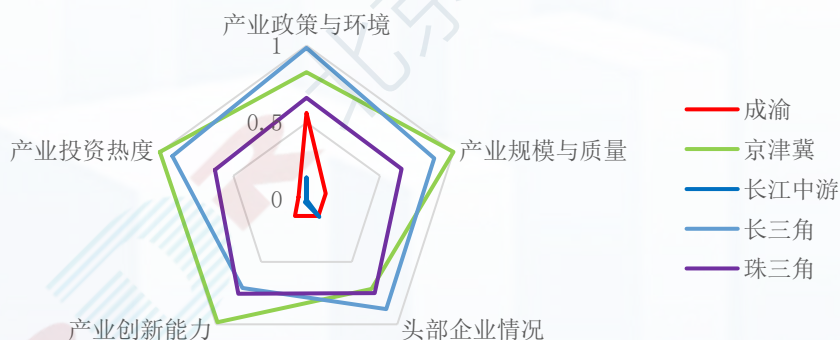
(三) 大数据产业发展区域对比

长三角、珠三角、京津冀领跑全国，整体格局基本稳定。

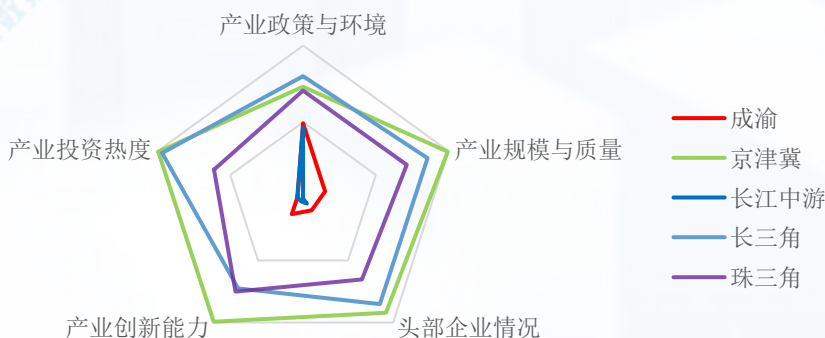
对需要优化提升的五大城市群分析来看，京津冀、长三角与珠三角发展较为全面，是大数据产业发展的核心区域，京津冀与长三角总指数位于第一名和第二名，珠三角在产业创新能力方面仅次于京津冀位于城市群第二位，成渝城市群、长江中游城市群在产业政策和环境方面也有突出表现。以北京为首的京津冀城市群头部企业情况排名明显提高，独角兽、瞪羚、高新技术企业在一年内发展较为迅速，同时产业规模与质量、产业创新能力以及产业投资热度几项指标都遥遥领先。

区域产业发展指数对比

2019各城市群整体对比



2020各城市群整体对比

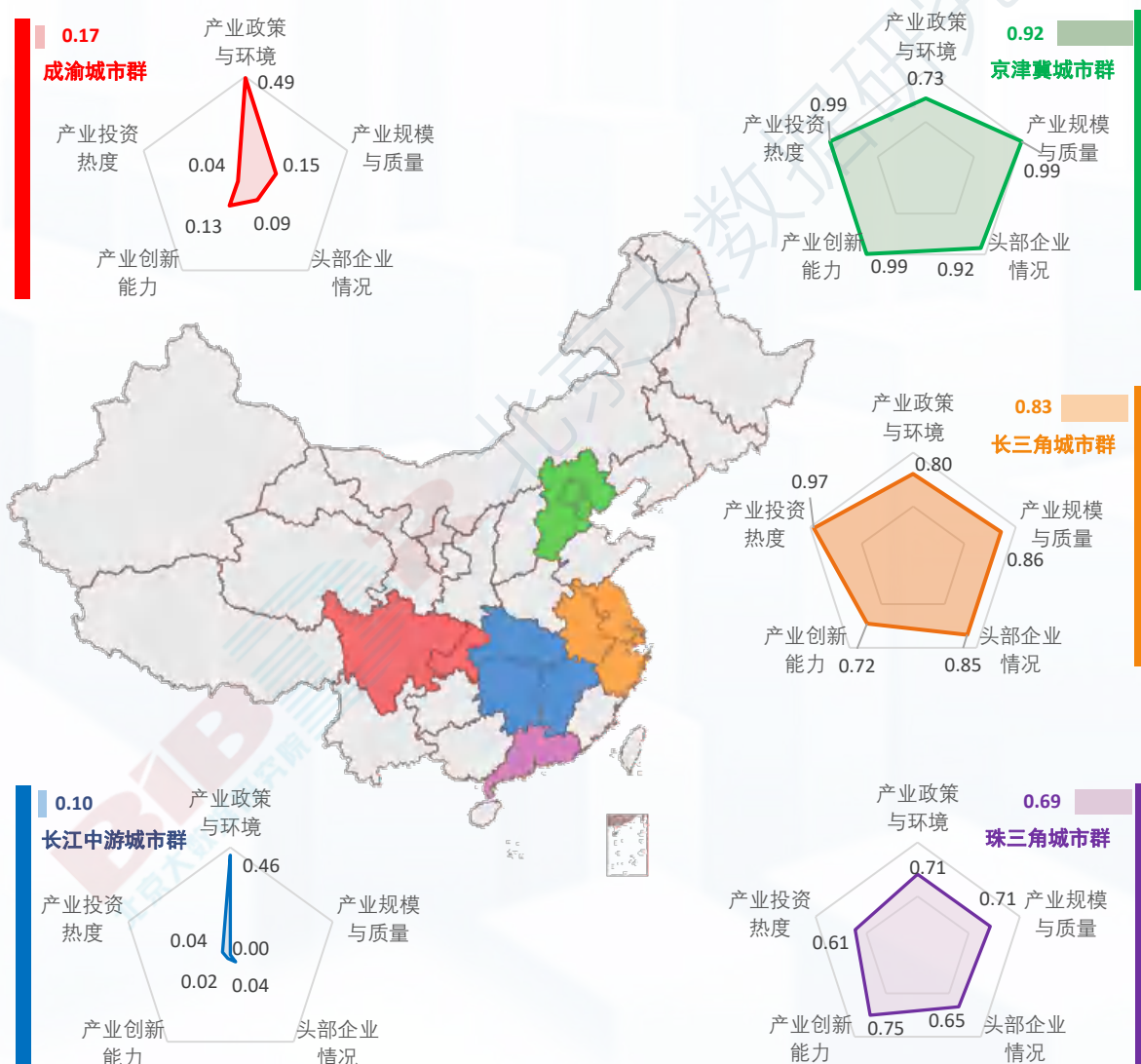


二、大数据产业发展指数总体评价

(三) 大数据产业发展区域对比

分别看五大城市群的大数据产业发展情况，京津冀、长三角与珠三角的整体较为均衡，但珠三角城市群的综合表现明显弱于长三角与京津冀两大城市群。长江中游城市群与成渝城市群情况相似，产业政策与环境的表现比其他一级指标表现更好。

区域产业发展指数对比



三、 大数据产业发展分指数评析



三、大数据产业发展分指数评析

一级指标排名变化情况（前十名）

总体来说，2020年大数据产业政策与环境、产业投资热度位次变化较大，而产业规模与质量、头部企业情况变化较小。

产业政策与环境



产业规模与质量



头部企业情况



产业创新能力



产业投资热度



三、 大数据产业发展分指数评析

(一) 产业政策与环境

管理机构基本完备，产业联盟和研究机构成为竞争热点。

2020年，随着大数据行业各项政策法规的完善和各地智慧城市战略的推进，大数据行业进一步规范、大数据应用进一步深入到公共治理当中，使得产业政策与环境成为影响大数据产业格局的重要因素。2020年，重庆、贵州、沈阳、内蒙古作为国家大数据综合试验区，继续保持政策优势，“大数据管理局”持续在产业发展中发挥着引领作用。2020年，135个城市没有新设立大数据主管机构，在国家层面也没有批复新的大数据综合试验区。从总体战略布局来看，大数据产业格局基本稳定，正逐渐进入细分领域的深化阶段。

产业政策与环境指数排名（前二十名）

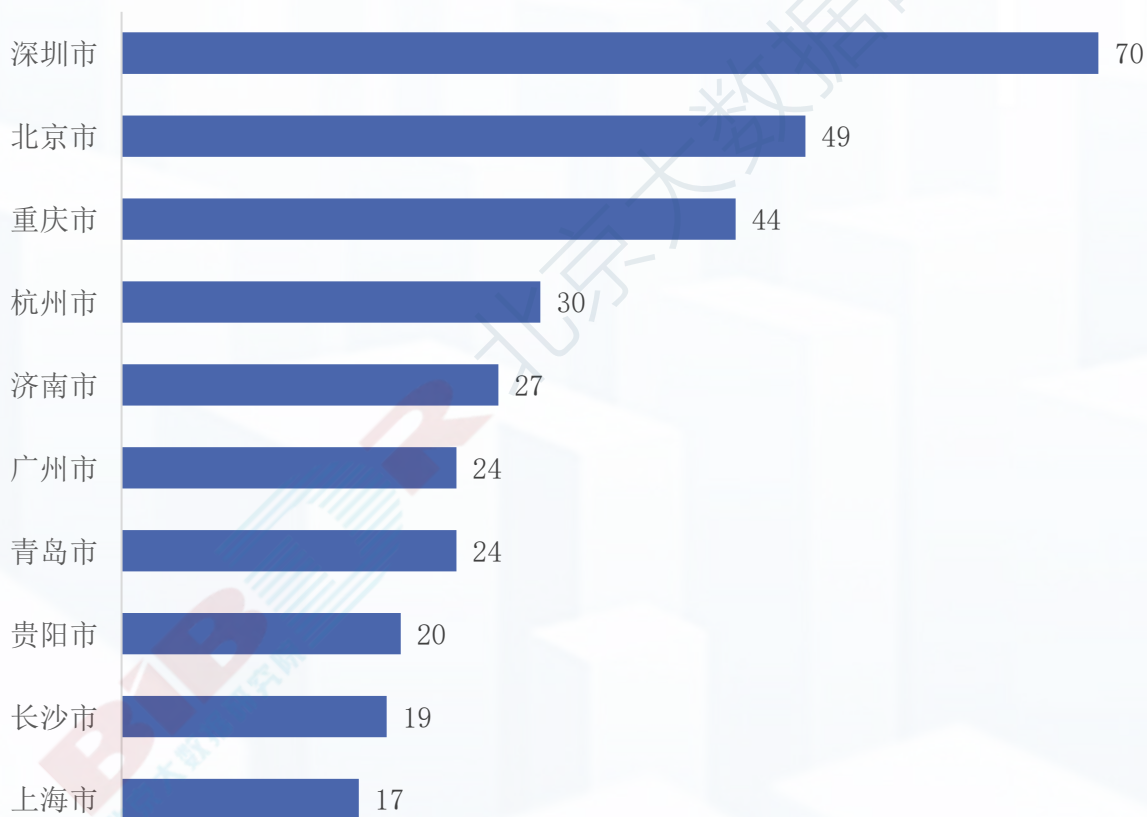
城市	2021 指数得分	2021 指数排名	2020 指数排名	城市	2021 指数得分	2021 指数排名	2020 指数排名
深圳市	0.98060	1	5	佛山市	0.77067	11	9
北京市	0.96940	2	1	惠州市	0.65141	12	10
重庆市	0.96520	3	2	杭州市	0.59583	13	23
广州市	0.90872	4	6	福州市	0.56254	14	22
上海市	0.90719	5	4	无锡市	0.54643	15	56
贵阳市	0.89954	6	3	济南市	0.53872	16	16
天津市	0.82814	7	7	长沙市	0.53868	17	25
呼和浩特市	0.79603	8	8	苏州市	0.53736	18	20
珠海市	0.79494	9	11	合肥市	0.53343	19	44
沈阳市	0.78832	10	12	青岛市	0.53113	20	18

三、 大数据产业发展分指数评析

(一) 产业政策与环境

在产业联盟和研究机构方面，各城市之间的差异显著。深圳市、重庆市产业联盟和研究机构数量迅猛增长，新增个数超过35家。大数据产业相对薄弱的城市也开始大规模设立相关机构，其中既包括交流性质的人工智能协会、商会，如福建省互联网经济促进会；也包括实体企业创立的大数据研究院，如海尔成立的青岛市数字家庭产业与应用促进会等。

2020年主要城市大数据产业联盟、研究机构数量（前十名）



三、 大数据产业发展分指数评析

(二) 产业规模与质量

疫情之下大数据产业依然稳中有增，北深上杭持续领跑全国。

2020年我国重要城市大数据产业规模与质量持续向好，总体排名保持稳定，前10名没有发生变化，各地大数据产业规模在疫情当中整体保持了稳步增长的态势。

产业规模与质量指数排名（前二十名）

城市	2021 指数得分	2021 指数排名	2020 指数排名	城市	2021 指数得分	2021 指数排名	2020 指数排名
北京市	1.00000	1	1	天津市	0.60515	11	11
深圳市	0.85179	2	3	厦门市	0.57781	12	12
上海市	0.83931	3	2	青岛市	0.57008	13	20
杭州市	0.80715	4	4	郑州市	0.55488	14	16
广州市	0.73621	5	5	福州市	0.53876	15	13
成都市	0.67955	6	6	无锡市	0.52847	16	17
南京市	0.67890	7	7	西安市	0.52049	17	14
武汉市	0.64738	8	8	重庆市	0.51832	18	15
苏州市	0.62480	9	9	长沙市	0.50419	19	19
合肥市	0.61660	10	10	济南市	0.48990	20	18

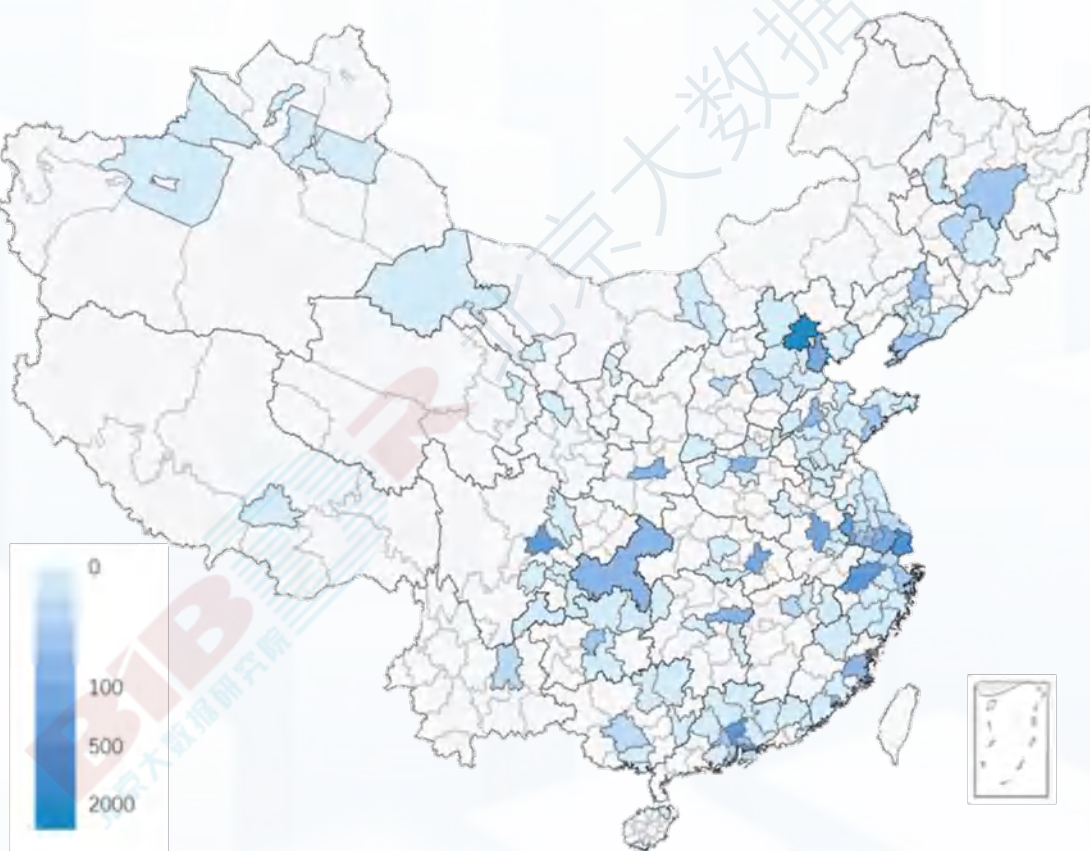
三、 大数据产业发展分指数评析

(二) 产业规模与质量

从企业分布来看，大数据企业主要分布在直辖市和东部沿海城市，呈现出高度的聚集性。

共有11座城市的企業数量超过100家，其中北京以2191家企业高居榜首，位列第二的上海有819家，第三名深圳市有635家。

全国大数据企业数量分布



三、大数据产业发展分指数评析

(二) 产业规模与质量

从人员规模来看，北京、深圳、杭州、青岛4座城市的大数据产业从业人员超过了10万人，其中北京以48.9万从业人员位居榜首，深圳以43万人紧随北京之后，深圳以较少的企业数量聚集了众多的从业人员，说明华为、腾讯等大型企业的头部效应较为显著。从企业平均员工人数来看，各城市间呈现出的差异较大，深圳市企业平均员工人数677人，杭州市企业平均员工人数350人、北京市企业平均员工人数233人、广州市企业平均员工人数102人，在一定程度上反映了各地大数据企业是以龙头企业为主，还是以小微企业为主。

大数据企业数量和人员规模情况（前十名）



三、大数据产业发展分指数评析

(三) 头部企业情况

北京头部企业表现超越深圳，在各类优质企业数量方面一骑绝尘。

头部企业情况呈现出鲜明的梯度分布，其中北京、深圳、杭州、上海4座城市是众多企业总部、投资基金聚集地和交易所驻地，得分在0.68到0.90之间，属于第一梯队。广州、南京、成都、无锡在0.42到0.53之间，属于第二梯队。其余城市得分在0.38及以下，分布较为密集，属于第三梯队。

头部企业情况指数排名（前二十名）

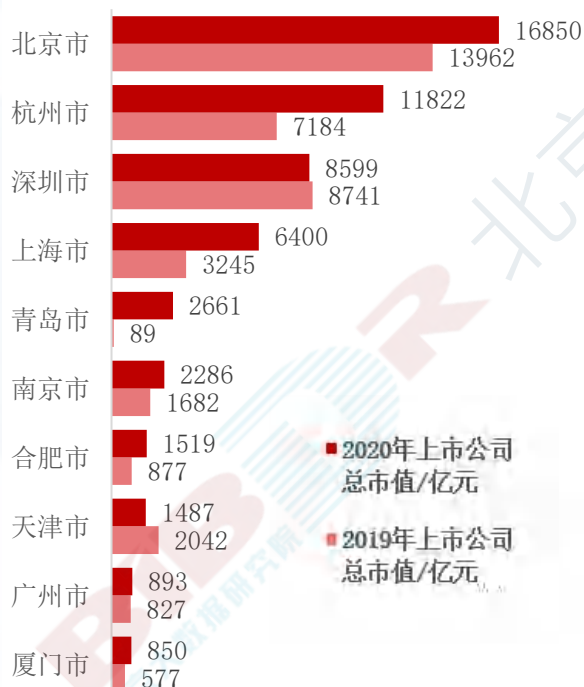
城市	2021 指数得分	2021 指数排名	2020 指数排名	城市	2021 指数得分	2021 指数排名	2020 指数排名
北京市	0.89778	1	2	济南市	0.36970	11	20
深圳市	0.74946	2	1	重庆市	0.36548	12	16
杭州市	0.72959	3	3	武汉市	0.35595	13	7
上海市	0.68646	4	4	苏州市	0.35240	14	11
广州市	0.52626	5	5	天津市	0.34093	15	9
南京市	0.50776	6	6	郑州市	0.33079	16	12
成都市	0.46089	7	8	厦门市	0.32267	17	13
无锡市	0.42154	8	10	青岛市	0.31821	18	21
合肥市	0.38113	9	14	长春市	0.31176	19	32
西安市	0.37842	10	15	福州市	0.31050	20	17

三、 大数据产业发展分指数评析

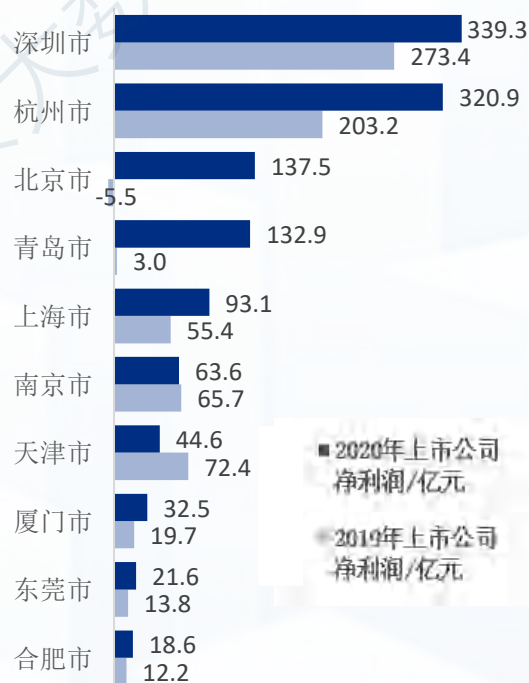
(三) 头部企业情况

在上市公司市值方面，北京以1.685万亿高居榜首，上海和杭州增长迅速，增幅达49%和39%。在净利润方面，深圳和杭州以339亿和321亿明显领先于其它城市，上升势头非常强劲。2020年青州市大数据企业在市值和净利润方面都出现了可观的增长，海尔智家实现盈利88.77亿元，青岛港5G智慧码头建设稳步推进，在“大数据+制造”“大数据+物流”方面的探索值得关注。

上市公司市值情况（前十名）



上市公司净利润（前十名）



三、 大数据产业发展分指数评析

(三) 头部企业情况

北京市在上市公司、独角兽企业、瞪羚企业和高新技术企业数量方面均高居榜首。除北京之外，大数据领域的独角兽企业主要分布在深圳（15家）、上海（11家）和杭州（9家），与综合指数领先城市情况基本一致。瞪羚企业方面，随着2020年8月全国首个瞪羚企业高端服务平台“光谷瞪羚源”在武汉成立，吸引众多瞪羚企业落户武汉光谷，截至目前武汉市有瞪羚企业86家，仅次于北京、广州、上海和杭州，位列全国第5。

上市企业、独角兽企业、瞪羚企业和高新技术企业分布



三、大数据产业发展分指数评析

(四) 产业创新能力

疫情冲击大数据企业创新能力，头部城市技术地位长期积淀难以撼动。

2020年，突如其来的新冠肺炎疫情限制了人才流动、迟滞了技术交流，一定程度上限制了大数据产业的发展进程。但综合两年排名情况看，排名变化很小，前12名均为一线城市或东部沿海城市，北京、深圳、上海、广州在产业创新能力方面遥遥领先，作为四大科教城市之一的合肥，以及积极推动建设国家科学中心的南京紧随其后，而13-20名中也不乏有郑州、长沙、成都、沈阳这样的中西部及东北城市。

产业创新能力指数（前二十名）

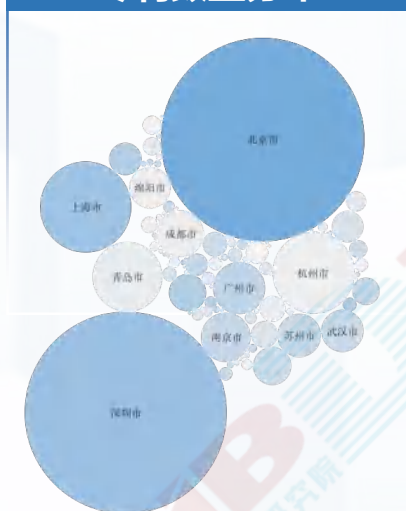
城市	2021 指数得分	2021 指数排名	2020 指数排名	城市	2021 指数得分	2021 指数排名	2020 指数排名
北京市	1.00000	1	1	福州市	0.43317	11	11
深圳市	0.70202	2	2	无锡市	0.42654	12	12
上海市	0.67210	3	3	郑州市	0.42055	13	13
广州市	0.53733	4	4	天津市	0.39996	14	14
合肥市	0.50031	5	5	芜湖市	0.39799	15	17
南京市	0.48416	6	7	长沙市	0.39612	16	16
厦门市	0.47900	7	6	成都市	0.38753	17	15
武汉市	0.46582	8	9	嘉兴市	0.38524	18	21
苏州市	0.46532	9	8	沈阳市	0.36503	19	18
杭州市	0.43929	10	10	常州市	0.35546	20	23

三、 大数据产业发展分指数评析

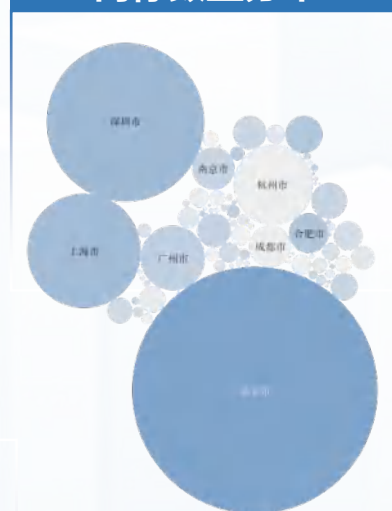
(四) 产业创新能力

在研发能力上，2020年受疫情影响各主要指标虽呈现出不同程度的下降，但北上广深4个一线城市以及江浙沪部分城市仍具有较强优势，其中北京、广东各项研发能力较为突出。R&D经费投入方面，北京市以6.31%明显领先，深圳、上海、合肥、厦门等位列其后，R&D强度达3%以上。在大数据企业软件著作权、专利、商标数量方面，北京、深圳、上海、杭州均排在前列。软件著作权数量方面北京一家独大，但在专利数量方面深圳与北京更为接近。表明大数据产业创新能力所依托的产业环境、人才环境、科研环境是长期积淀的结果，头部城市的技术地位稳固。

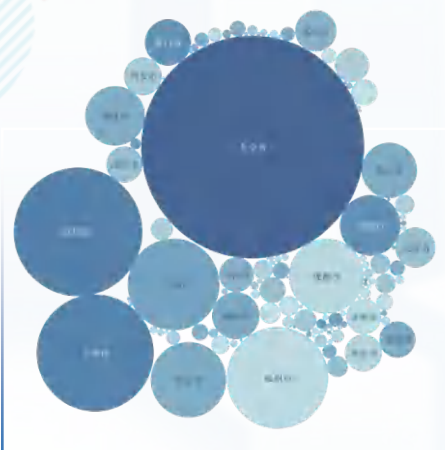
专利数量分布



商标数量分布



软件著作权数量分布



三、 大数据产业发展分指数评析

(五) 产业投资热度

北京继续领跑、杭州势头迅猛，二三线城市中显露新秀。

在产业投资热度方面，北京高居榜首，杭州超过上海位列第二，广州增幅较小，深圳有所下滑，其余城市得分相近。值得注意的是，2020年嘉兴、宁波、上饶3座城市的投资热度增长迅速，一举进入前20名，显示长三角地区大数据产业投资有向二线、三线城市延伸的态势。同时，珠海、哈尔滨等部分二三线城市在其它分指数榜单中优势并不明显，但也具有相当的投资热度。

产业投资热度指数（前二十名）

城市	2021 指数得分	2021 指数排名	2020 指数排名	城市	2021 指数得分	2021 指数排名	2020 指数排名
北京市	1.00000	1	1	嘉兴市	0.45365	11	46
杭州市	0.91851	2	3	天津市	0.41441	12	6
上海市	0.82866	3	2	合肥市	0.38997	13	11
深圳市	0.77941	4	4	厦门市	0.37964	14	15
广州市	0.69561	5	5	宁波市	0.35114	15	25
苏州市	0.59530	6	8	上饶市	0.35108	16	27
成都市	0.57017	7	9	贵阳市	0.34227	17	17
南京市	0.53016	8	7	福州市	0.33811	18	16
武汉市	0.48854	9	10	珠海市	0.32988	19	19
无锡市	0.46563	10	12	哈尔滨市	0.31158	20	28

三、 大数据产业发展分指数评析

(五) 产业投资热度

自2018年“新基建”概念提出以来，工业互联网逐渐成为大数据产业发展的增长点。2020年，工业和信息化部发布了《关于推动工业互联网加快发展的通知》和《关于工业大数据发展的指导意见》，指出要发挥工业互联网在产业链协同、货物标识体系、安全保障体系、复工复产等方面的应用。工业互联网的发展将促使积聚在北上广深杭的大数据人才与技术外溢到更多二三线城市，带来大数据产业投资的“下沉”浪潮。

2020年全国大数据企业融资总额远超1万亿元。大数据产业的投资分布呈现出鲜明的“二八效应”，其中北京占据融资总额约一半、杭州占据融资总额约1/4，在投资热度方面领先全国。

大数据企业融资额热力图



四、 展望与建议



四、 展望与建议

(一) 发挥数据资源禀赋作用，促进区域协同发展

围绕数据资源，整合技术、产品、服务等全产业链条中的主体，发挥数据资源禀赋作用，多角度、多层面促进大数据产业发展，搭建产业生态，打造以数据带动产业发展，以产业推动区域建设的路径，促进形成区域协同发展的新格局。

(二) 聚焦数据立法，进一步加强行业自律

随着《数据安全法》《网络安全审查办法》等一系列政策法规的出台，对大数据产业在保障数据安全、规范数据使用上提出了明确的要求。大数据企业应进一步加强行业自律，在数据收集、使用、共享、交易等过程中自觉遵守法律规定，杜绝出现“灰色地带”，切实保障数据安全。

四、 展望与建议

(三) 打造数据发展新模式，提高产业创新能力

进一步优化大数据产业发展环境，整合协调政府、市场、学校、研究机构等多方资源，推动“政产学研用”协同发展，积极探索创新模式，有效激发科技活力，提高大数据产业创新能力。

(四) 深化数据融合应用，探索数据要素市场化发展路径

紧跟国家战略，深度融合不同行业的数据，有效挖掘数据价值，探索数据要素市场化的发展路径，聚焦数据本质，应用下沉，做实大数据产业，释放数据价值，彰显数据在经济发展中的作用。

编制单位介绍

北京大数据研究院成立于2015年，在北京市委市政府指导下，由中关村管委会、海淀区政府、北京大学、北京工业大学四方共同支持建立，是大数据和人工智能领域第一个集人才培养、科研创新、产业转化以及智库咨询为一体的产学研综合性平台。

大数据分析与应用技术国家工程实验室由国家发展和改革委员会批复建设，由北京大学牵头，以北京大数据研究院为重要支撑实施单位，联合中国科学院数学与系统科学研究院、中山大学、北京奇虎科技有限公司、北京嘀嘀无限科技发展有限公司、中国信息安全研究院有限公司等多家单位共建，于2017年3月21日在北京大学正式揭牌。实验室的主要建设内容是在现有研发和试验条件基础上，建设大数据分析系统开发、可视化展示、测试与评估、重大应用示范与系统集成研发平台。

大数据分析技术创新中心是北京大数据研究院和大数据分析与应用技术国家工程实验室联合共建的产业化合作与服务的重要载体与平台。中心以城市治理与决策的数字化与智能化转型赋能为目标，致力于打造智慧城市大数据应用创新方案与产品的重要策源地与示范推广平台。

中心建立了企业数据库和政策法规库，已发布《京津冀大数据产业地图》、《大数据产业发展指数》、《典型城市大数据发展报告》、《数字经济产业发展指数报告》等多项系列成果，为各级政府部门和企业提供大数据顶层设计、大数据平台建设、大数据标准规范制定、数据资源管理、数字经济发展规划、大数据产业园区规划、智慧城市建设等咨询服务；并通过自然语言处理、知识图谱、机器学习等大数据分析技术和方法，为政府部门和企业提供基于数据的决策和应用落地服务，支撑政府决策和企业管理。

编写组成员

指数负责人：程超

指数专家团队：贾金柱 傅毅明 王娟

指数团队成员：陈杨 冷莹 蔡婕
刘志攀 贾雅洁 刘兰馨
李枫秋 焦之贤 吴燕妮
金泊翰 郭杭

联系人：程超

邮箱：chaoc@bibdr.org



扫码关注
北京大数据研究院



扫码关注
大数据分析技术创新中心