



2021 数字中国创新大赛大数据赛道-城市管理大数据 专题

——操作手册

1 登录

登录 data.xm.gov.cn，点击赛事详情进入“2021 数字中国创新大赛大数据赛道-城市管理大数据专题”网站首页。点击右上角“登录”按钮或“报名参赛”跳转至登录界面，可使用 i 厦门个人实名认证登录或微信扫码登录，如图所示：



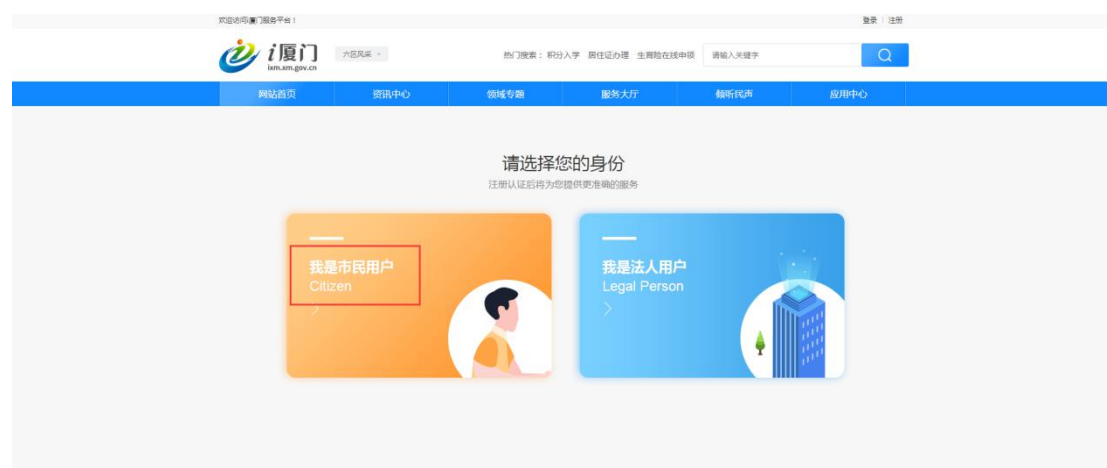
1.1 i 厦门实名认证注册流程

已有 i 厦门个人实名账号的用户,输入账号密码登录系统即可进入赛事系统。

未有 i 厦门个人实名账号的用户, 点击“立即注册”, 跳转至 i 厦门注册页面进行账号注册, 如图所示:

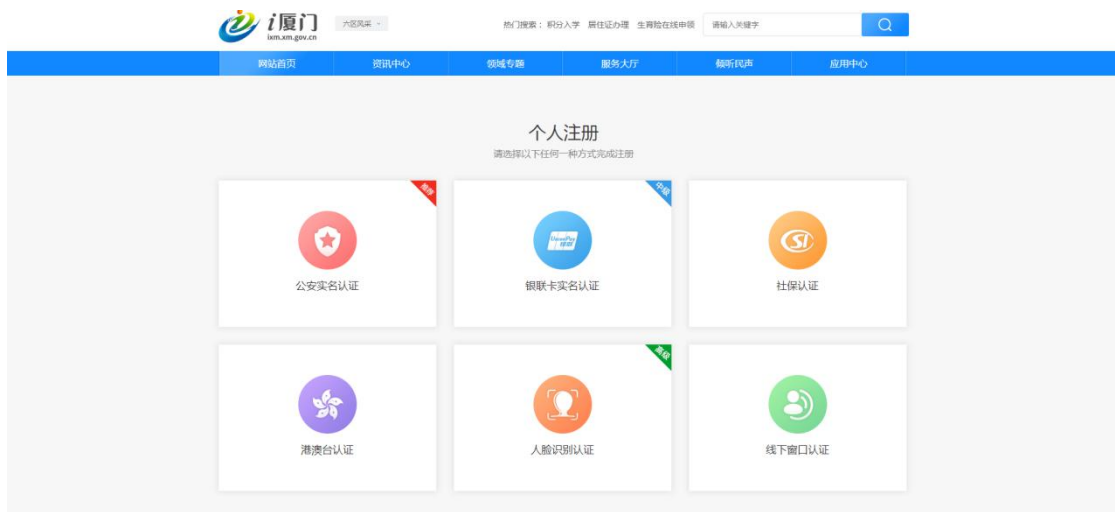


进入 i 厦门注册页面, 勾选“我是市民用户”, (“2021 数字中国创新大赛大数据赛道-城市管理大数据专题”仅支持 i 厦门个人实名账号登录), 如图所示:



根据个人情况, 选择对应的注册方式, 以下以公安实名认证流程为例, 若公安实名认证失败, 可尝试银联卡实名认证等其他注册方式。注册完毕后, 即可使用账号密码登录“2021 数字中国创新大赛大数据赛道-城市管理大数据专题”网

站，如图所示：



根据勾选的注册方式进入对应的注册页面，根据提示填写真实姓名和身份证号，点击下一步，如图所示：



填写手机号和收到的验证码，**验证码收不到的情况下，请多点击几次“获取验证码”**，点击下一步，如图所示：

公安实名认证

1 身份认证

2 短信验证

3 设置密码

请输入您的真实信息，通过后则不能更改

手机号：

157

图片校验码：

验证成功

验证码：

请输入验证码

167秒后重新发送

下一步

填写密码，点击注册，如图所示：

公安实名认证

1 身份认证

2 短信验证

3 设置密码

请输入您的真实信息，通过后则不能更改

设置密码：

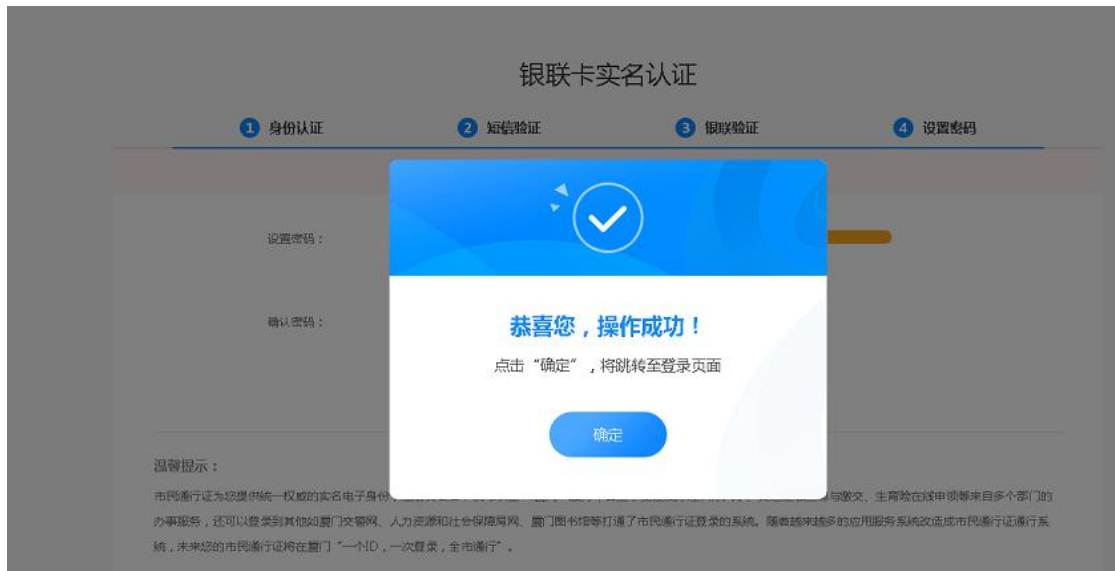
密码长度8-30位要包含字符和数字，可以使用特殊字符

确认密码：

*****|

注册

给出提示：恭喜您，操作成功！点击确定可回到登录页面，使用账号密码即可进入“2021 数字中国创新大赛大数据赛道-城市管理大数据专题”网站，如图所示：



2 赛题报名

2.1 创建队伍

未登录的情况下，请先登录赛事系统。已登录的情况下，点击【首页】“报名参赛”或进入【赛题说明】的赛题详细根据情况对赛题进行报名，在报名界面勾选“自己创建队伍（同样适用于个人参赛）”即可创建队伍。队长可在【我的比赛-我的队伍-组队通知】中查看队员提交的组队申请，如图所示：



① 请注意，自己创建队伍或成功加入已有队伍后，将无法加入其他队伍。

* 选择组队方式

☒ 自己创建队伍（同样适用于个人参赛）

☐ 加入已有队伍

* 队伍信息

队伍名称

请输入队伍名称

队伍类型

请选择队伍类型

参赛单位/学校名称

请输入学校或单位名称，可填“无”，多个单位可用逗号，隔开

2.2 加入已有队伍

用户要加入已有队伍时，**未登录的情况下，请先登录赛事系统**。已登录的情况下，点击【首页】“**报名参赛**”或进入【赛题说明】的赛题详细根据情况对赛题进行报名，勾选“**加入已有队伍**”后，搜索要加入的队伍名称后，点击“提交申请”即可向队长发起组队申请，队长审核通过后，在【我的比赛】-【我的队伍】中可查看到队伍信息和提交的申请情况。如图所示：



ⓘ 请注意，自己创建队伍或成功加入已有队伍后，将无法加入其他队伍。

* 选择组队方式

☐ 自己创建队伍（同样适用于个人参赛） ☒ 加入已有队伍

* 队伍信息

请输入队伍名称搜索 ▼

* 个人邮箱

wang .com.cn

3 队伍信息确认/实名

3.1 队伍成员实名

参赛者参赛报名后，如需下载赛题数据需完成个人实名认证方可下载；如需提交作品需参赛队伍下的全部队员完成个人实名认证方可提交作品。

未实名的队员需用 i 厦门个人实名账号登录“2021 数字中国创新大赛大数据赛道-城市管理大数据专题”网站，完成个人实名认证操作。

若队长需修改队伍信息或新增队员需在报名截止前完成，报名截止后将不允许再变更队伍信息或新增成员。

3.2 队伍信息确认

提交作品前，队长需进入【我的比赛-我的赛题-提交作品】进行队伍信息的确认。

若队伍成员未全员实名，则需全员实名后才可进行确认操作。若队伍成员已全员实名，则点击“去确认”按钮对队伍信息进行确认，如下图所示。

队伍信息一旦确认后，**不可增减队伍成员和修改队伍信息**，请谨慎操作。



4 作品提交账号获取

队伍成员信息确认后，进入【我的比赛-我的赛题-提交作品】则可获取作品提交地址、登录账号、密码以及大赛操作手册，如下图所示。



5 作品提交

本次赛事作品提交是通过 SFTP 方式来完成的。SFTP 登陆方式可通过命令行方式或 SFTP 工具来完成，每支队伍的 SFTP 空间大小是 5G，请控制好自己提交的文件大小。SFTP 账号密码为官方下发的账号密码，详见本文档第 4 章节-作品提交账户获取。【注】：sftp 端口是 57891。

5.1 命令行方式登录 SFTP

可通过 filezilla 或其它 sftp 工具，连接登陆，参照如下：

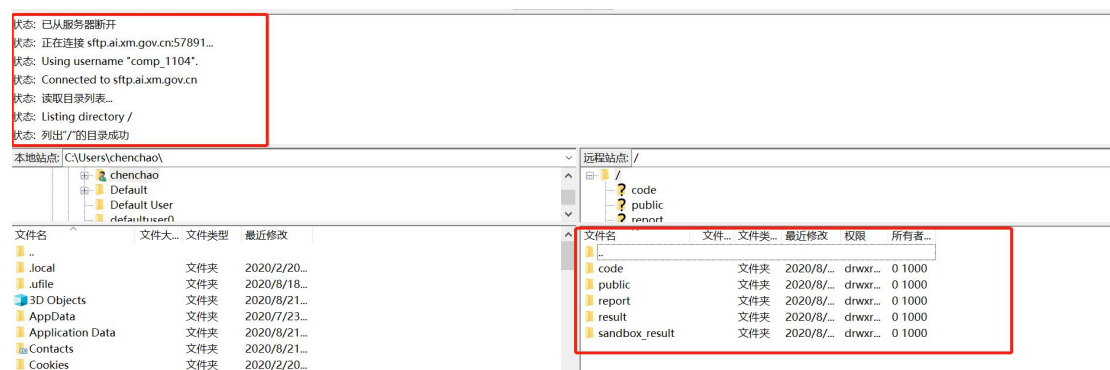
```
[root@172.16.1.100 ~]# sftp -oPort=57891 comp_882@sftp.ai.xm.gov.cn
comp_882@sftp.ai.xm.gov.cn's password:
Connected to sftp.ai.xm.gov.cn.
sftp> ls
code                public              report              result              sandbox_result
sftp>
```

备注：如上命令中 comp_882 为选手比赛账号示例，实际操作中选手需要更换为自己的比赛账号。

5.2 SFTP 工具方式登陆 SFTP

可通过 filezilla 或其它 sftp 工具连接登陆，在【文件】—【站点管理器】里设置服务端的主机、端口，参照如下：





备注：如上命令中 `comp_1104` 为选手比赛账号示例，实际操作中选手需要更换为自己的比赛账号。

5.3 作品提交所在文件路径

本次大赛，参赛队伍需提交的作品包含：①算法模型分析结果集、②算法模型源码、③赛题分析报告、PPT 答辩材料及其他自主选择提交的交付材料。对应 SFTP 指定目录如下表：

序号	作品名称	存放目录	命名规则
1	早高峰共享单车潮汐点的群智优化赛题对应的算法模型源码、算法模型等	/code/gxdc	code-参赛编号-提交日期
2	城市早晚高峰时段综合运力智能调度应用赛题对应的算法模型源码、算法模型等	/code/zwgf	code-参赛编号-提交日期
3	城市绿色慢行交通的友好度评价研究赛题对应的算法模型源码、算法模型等	/code/lsmx	code-参赛编号-提交日期
4	早高峰共享单车潮汐点的群智优化赛题对应的分析报告、PPT 答辩材料及其他自主选择提交的交付材料	/report/gxdc	report-参赛编号-提交日期
5	城市早晚高峰时段综合运力智能调度应用赛题对应的分析报告、PPT 答辩材料及其他自主选择提交的交付材料	/report/zwgf	report-参赛编号-提交日期
6	城市绿色慢行交通的友好度评价研究赛题对应的分析报告、PPT 答辩材料及其他自主选择提交的交付材料	/report/lsmx	report-参赛编号-提交日期
7	早高峰共享单车潮汐点的群智优化赛题- 算法模型分析结果集	/result	result.txt

6 算法模型代码提交说明

参赛队伍提交的算法应为可稳定运行的算法模型，根据赛题规则要求，规范化模型输出，并整理完整代码、注释说明及补充材料，以备代码复核验证。

代码 (含源代码、算法模型等，不含算法输出结果、分析报告等)提交要固

定输出到 SFTP 的 **/code/[对应赛题英文缩写]** 目录下，如果选手参与多个赛题，则每个赛题对应一份算法模型代码。

注意：各文件命名要清晰罗列，提交时建议文件压缩后上传，文件命名为“**code-参赛编号-提交日期**”。

7 报告及答辩材料提交说明

参赛队伍提交的作品(含分析报告、答辩 PPT 材料等，不含算法结果文件、源代码、算法模型)提交要固定输出到 SFTP 的 **/report/[对应赛题英文缩写]** 目录下，如果选手参与多个赛题，则每个赛题对应一份分析报告和 PPT 答辩材料。

注意：各文件命名要清晰罗列，作品如包含多份材料统一打包成一份文件夹，文件命名格式为“**report-参赛编号-提交日期**”。分析报告需在 **2021 年 3 月 24 日 24 时**前提交，逾期提交则视为无效提交。

8 选手算法输出结果及提交说明

8.1 赛题名称

早高峰共享单车潮汐点的群智优化

8.2 赛题说明

共享单车，延伸了城市公共交通脉络，解决了市民出行“最后一公里”问题。然而，随着共享经济模式被越来越多市民接受，成为出行习惯，潮汐现象也随之出现。白天工作、晚上休息的人类活动规律的客观存在，加之上下班时间段的集中，导致早晚高峰“一车难寻”、“无地可停”的供需矛盾。本题希望通过对车

辆数据的综合分析，对厦门岛内早高峰阶段潮汐点进行有效定位，进一步设计高峰期群智优化方案，缓解潮汐点供需问题，以期为城市管理部门和共享单车运营方研究制定下一步优化措施提供数据支撑。

8.3 赛题任务

任务一：为更好地掌握早高峰潮汐现象的变化规律与趋势，参赛者需基于主办方提供的数据进行数据分析和计算模型构建等工作，识别出工作日早高峰 07:00-09:00 潮汐现象最突出的 40 个区域，列出各区域所包含的共享单车停车点位编号名称，并提供计算方法说明及计算模型，为下一步优化措施提供辅助支撑。

任务二：参赛者根据任务一 Top40 区域计算结果进一步设计高峰期共享单车潮汐点优化方案，通过主动引导停车用户到邻近停车点位停车，进行削峰填谷，缓解潮汐点停车位（如地铁口）的拥堵问题。允许参赛者自带训练数据，但需在参赛作品中说明所自带数据的来源及使用方式，并保证其合法合规。（城市公共自行车从业者将发生在早晚高峰时段共享单车“借不到、还不进”的问题称之为“潮汐”现象。本题涉及的“潮汐现象”聚焦“还不进”的问题，识别出早高峰共享单车最淤积的 40 个区域）。

8.4 模型预测任务说明

为更好地掌握早高峰潮汐现象的变化规律与趋势，参赛者需基于主办方提供的数据进行数据分析和计算模型构建等工作，识别出工作日早高峰 07:00-09:00 潮汐现象最突出的 40 个区域，列出各区域所包含的共享单车停车点位编号名称，并提供计算方法说明及计算模型，为下一步优化措施提供辅助支撑，

城市公共自行车从业者将发生在早晚高峰时段共享单车“借不到、还不进”的问题称之为“潮汐”现象。本题涉及的“潮汐现象”聚焦“还不进”的问题，识别出早高峰共享单车最淤积的 40 个区域。

8.5 运行结果提交格式

算法运行结果要固定输出到 sftp 的 **/result** 目录下，结果文件为 **result.txt**，格式为 **utf-8**。比赛平台最终会去取这个结果文件算出结果，如果提交结果为非 result.txt 文件命名，或没有提交到/result 目录下，则不会有分数产生。算法运行结果文件内容中，**每个字段间以英文半角字符的竖线号 “|” 符分隔。**

选手提交的 FENCE_ID 与标准答案中 FENCE_ID 完全匹配。

注：

- 停车点位数据中共有 14071 个不同名称的电子围栏编号（即 FENCE_ID 共 14071 个），选手提交的答案需要列出所有的 FENCE_ID 对应的 FENCE_TYPE。
- 选手需至少提交 FENCE_ID、FENCE_TYPE 两个字段的值，才能进行评分。其中,提交的 FENCE_ID 值与个数必须与停车点位数据中的 FENCE_ID 相同，FENCE_TYPE 数值不能为空或范围外数值。

8.6 输出结果字段说明

➤ 选手算法输出结果格式样例：

FENCE_ID | FENCE_TYPE | BELONG_AREA

长青路_R_9 | 1 | XXXX

长青路_R_10 | 0 | XXXX

输出结果字段注释		
字段名	字段释意	备注
FENCE_ID	电子围栏唯一编号	{长青路_R_9, 长青路_R_10, ...}
FENCE_TYPE	是否属于在 40 个潮汐现象最明	{0: 否; 1: 是}

	显的区域内	
BELONG_AREA	所属区域	FENCE_TYPE 值为 1 时，电子围栏所对应的所属区域的名称（40 个）

注：输出结果文件里需要有表头，结果文件为 Utf-8 格式的 result.txt 文件，文件内每字段以**英文半角竖线**“|”间隔。

➤ 评测计分规则

根据由我们提供的标准答案（FENCE_ID、FENCE_TYPE），计算 Precision 值、Recall 值以及 F1-score 值，下面为计算步骤：

1) 混淆矩阵（Confusion Matrix）

首先通过标准答案（真实值）与选手提交的结果（分析值），根据 FENCE_ID 与 FENCE_TYPE 的分布分别计算各个类别的个数。其中：

- 真实值为 0，分析值为 0：真实值不在潮汐范围内，分析值也不在潮汐范围内。统计出来的结果标为 TN。
- 真实值为 0，分析值为 1：真实值不在潮汐范围内，分析值在潮汐范围内。统计出来的结果标为 FP。
- 真实值为 1，分析值为 0：真实值在潮汐范围内，分析值不在潮汐范围内。统计出来的结果标为 FN。
- 真实值为 1，分析值为 1：真实值在潮汐范围内，分析值也在潮汐范围内。统计出来的结果标为 TP。

表 1 混淆矩阵

(分析值) (真实值)	0	1	
0	TN	FP	TN+FP

1	FN	TP	TP+FN
	TN+FN	TP+FP	

2) 计算 Precision 与 Recall

选手所给的分析值与真实值都不在潮汐范围内（也就是 FENCE_TYPE 都为 0 的）：

$$\text{Precision} = \text{TN}/(\text{TN}+\text{FN})$$

$$\text{Recall} = \text{TN}/(\text{TN}+\text{FP})$$

选手所给的分析值与真实值都在潮汐范围内（也就是 FENCE_TYPE 都为 1 的）：

$$\text{Precision} = \text{TP}/(\text{TP}+\text{FP})$$

$$\text{Recall} = \text{TP}/(\text{TP}+\text{FN})$$

3) 计算 F1_Score

$$\text{F1_Score} = 2 * (\text{Precision} * \text{Recall}) / (\text{Precision} + \text{Recall})$$

选手分析值与真实值都是不在潮汐范围内（FENCE_TYPE 都为 0）：

$$\text{F0} = \text{F1_Score}(0)$$

选手分析值与真实值都是在潮汐范围内（FENCE_TYPE 都为 1 的）：

$$\text{F1} = \text{F1_Score}(1)$$

最后计算两类的平均值得出 $\text{F1_Score}(\text{avg}) = (\text{F0} + \text{F1})/2$ ，即 $\text{F1_Score}(\text{avg})$ 为选手提交结果的算法模型结果，再以 **30* F1_Score(avg)**作为算法题任务一的得分。

本题初赛成绩中，**任务一输出的算法模型得分占总得分 30%，任务二输出的优化方案占总得分 70%。**

8.7 算法平台成绩结果说明

选手提交的算法结果文件，平台大概约 1-2 小时周期评分一次，每个选手一天最多只评三次。如果选手一天评分超过三次，则当天无论再提交多少份结果文件都不再参与评分；如果选手提交的算法结果文件在同一个周期内多次提交，则只以最后一次提交的文件参与评分；最终算法分取选手历史分数最高的一次。

选手可登录 <https://data.xm.gov.cn/contest-series/digit-china-2021/#/3/> 具体查看每次算法得分结果及历史提交明细。

附录 1：常见问题 Q&A

1. Sftp 登陆时报 Permission Denied?

可能的原因有 2 种，一种是密码输出错误，请确保密码输入正确，可使用复制粘贴密码的方式，注意密码前后不要有空格。

另一种是当前 sftp 服务器正忙，可稍等一会重新提交。为保证选手能顺利通过 sftp 提交，避免尽量在比赛结束前的两三天内集中提交。

2. Sftp 上传文件后空间不足怎么办？

每个选手是 5G 的空间，选手尽量把作品等文件压缩打包后上传，如果压缩后空间还是满足不了，可联系官方助手(群里或电话)帮忙后台人工协助处理。

3. 任务一中提到的“识别出最突出的 40 个区域”中的“区域”具体指什么呢？

区域是选手根据自己的方法识别出的 07:00-09:00 潮汐现象最突出的区域所

在位置的名称，选手可以参考 FENCE_ID 属性值命名，也可参考地图上区域所在位置的名称来的，赛题对区域位置所在名称命名规则不做统一要求。